DOCUMENT RESUME

ED 318 398 IR 014 295

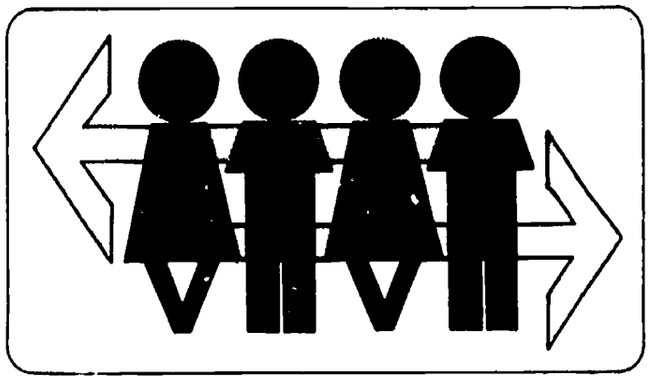
AUTHOR Baker, Robert F.; And Others
TITLE Building Capacity for Improvement of Educational Practice: An Evaluation of NIE's State Dissemination Grants Program. Volume IV: A Study of the Development of Scales Measuring Dissemination Capacity.
INSTITUTION NTS Research Corp., Durham, N.C.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Apr 81
CONTRACT 400-76-0166
NOTE 90p.; For volume I, see IR 014 293; for volume III, see IR 014 294; for volume V, see IR 014 296; volume II: 1979 State Abstracts is not available. For the 1978 state abstracts, see ED 178 099; for the 1979 interim report, see ED 184 532.
PUB TYPE Statistical Data (110) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Educational Improvement; Educational Practices; Elementary Secondary Education; Federal Programs; Improvement Programs; *Information Dissemination; *Scaling; State Departments of Education; Tables (Data)
IDENTIFIERS *Rasch Scaled Scores; State Capacity Building Program; *State Dissemination Grants Program

ABSTRACT

This report supplements the final evaluation of the State Dissemination Grants Program (SDGP), a major initiative within the mission of the National Institute of Education to assist state educational agencies in implementing, strengthening, and institutionalizing dissemination services that improve educational practice and equity. The introduction to this substudy explains that data collected for the final evaluation were subjected to a scaling analysis to: (1) confirm that indicants grouped together on the basis of their content were, in fact, measuring the same thing; (2) eliminate indicants not related to the other indicants in the facet; (3) parsimoniously summarize a large body of data; and (4) serve as a measuring device, by assigning facet scores to each scale for use in characterizing state capacity. The four main sections of this report discuss the five-component conceptual framework used to specify and organize the variables which define the process of building state dissemination capacity; the scaling procedures and the application of Rasch scaling analysis; the degree to which the scales provide a developmental perspective to the capacity building process; and interpretation and utilization of the resultant information for program improvement. Data are presented in 13 tables and 14 figures. (NRP)

Volume IV.
A Study of the Development of Scales Measuring Dissemination Capacity



THE STATE DISSEMINATION GRANTS PROGRAM

Building Capacity for Improvement of Educational Practice

Prepared for:

Research and Educational Practice Program
Dissemination and Improvement of Practice
National Institute of Education
Washington, D.C. 20208

Prepared by:
NTS RESEARCH CORPORATION
2634 Chapel Hill Blvd.
Durham, N.C. 27707

April, 1981

A Study of the Development of Scales Measuring
Dissemination Capacity

The report presented herein was prepared pursuant to Contract 400-76-0166 with the National Institute of Education, United States Department of Health, Education, and Welfare. Contractors undertaking such projects with government sponsorship are encouraged to express their professional judgment in the conduct of the project. Points of view or opinions do not, therefore, necessarily represent official positions or policies of the National Institute of Education.

BUILDING CAPACITY FOR IMPROVEMENT OF EDUCATIONAL PRACTICE: AN EVALUATION OF NIE'S STATE DISSEMINATION GRANTS PROGRAM

VOLUME IV: A STUDY OF THE DEVELOPMENT OF SCALES MEASURING DISSEMINATION CAPACITY

Robert F. Baker
Jolinda K. Decad
Eugene C. Royster
Doren L. Madey

Prepared for:

Research and Educational Practice Program
Dissemination and Improvement of Practice
National Institute of Education
Washington, D.C. 20208

Prepared by:

Educational Research Center
NTS Research Corporation
2634 Chapel Hill Blvd.
Durham, N. C. 27707

April 1981

PREFACE

The State Dissemination Grants Program is a major initiative within the mission of the National Institute of Education (NIE) "to promote educational equity and improve the quality of educational practice." NIE expects the State Dissemination Grants Program (SDGP) will aid the development of a nationwide capability for educational improvement by assisting a significant group of actors, state education agencies (SEAs), to implement, strengthen, and institutionalize dissemination services that improve educational practice and equity.

Under the sponsorship of the NIE's Program on Dissemination and Improvement of Practice, NTS Research Corporation conducted a multi-year study of the State Dissemination Grants Program (NIE Contract No. 400-76-0166, October 1976 - April 1980) to address two major questions:

- Is dissemination capacity being built as a result of this program? If so, how?
- Is the program having an effect? If so, what is the nature of the effect?

The evaluation was in two phases. Phase I of the study, an eleven-month design phase that extended from October 1976 through August 1977, was devoted to describing the program, clarifying and translating the program's goals into measurable variables, and developing a design, appropriate instrumentation, and data collection and analysis procedures for the study. Familiarization visits to 23 projects, refinements in the study design, and approval of a forms clearance package occurred during September 1977 - August 1978. Phase II, the full-scale evaluation, was initiated in September 1978 and concluded in April 1980. Phase II objectives included describing and tracking the process of building dissemination capacity, documenting the impact of the program, sharing the study findings and analyses with NIE and the states to promote program and project improvement, and developing mechanisms for the continual evaluation and measurement of dissemination capacity.

The final report for the NTS study is comprised of four volumes:

- This volume, Volume IV: A Study of the Development of Scales Measuring Dissemination Capacity (April 1981) is a technical report which describes how the scales were developed and how they have been used.

- Volume I: Building Capacity for Improvement of Education: An Evaluation of NIE's State Dissemination Grants Program (April 1981), is the final evaluation report of the State Dissemination Grants Program. Included are descriptions of the program and the evaluation, of qualitative cross-case analyses of five capacity building states, generic descriptions of state dissemination systems, and quantitative analyses which identify factors which facilitate and impede the development and institutionalization of SEA dissemination systems. The analyses reveal that dissemination capacity is being built, participation in the program enhances such growth, and SEA dissemination systems of states participating in the program differ from those of non-participating states. A final chapter discusses the policy implications of these and other findings.

- Volume II: 1979 State Abstracts: State Dissemination Efforts (April 1980), profiles dissemination activities in thirty-eight SEAs as of December 1979. In addition to summaries of capacity building project states, this document describes the status of dissemination efforts in states that have not participated in the program.

- Volume III: A Study of Linker Activities and Roles (April 1981), describes how people help others access and use information for school improvement. The study is based on data collected from linkers associated with the program.

Prior to 1980, seven major reports were prepared under Phase II of the NTS study:

- 1978 State Abstracts (March 1979) contains summaries of dissemination activities in twenty-nine SEAs as of November 1978. Included are nine SEAs initially funded in 1975, fourteen additional projects initially funded in 1976, and six SEAs initially funded in 1977. An introductory chapter presents an analysis across the individual projects.

- Building Capacity for Improvement of Education: An Evaluation of NIE's State Dissemination Grants Program, Interim Report, (July 1979) is the interim report on the full-scale evaluation of the State Dissemination Grants Program. Included is an overview of the evaluation, purposes of the study, framework, evaluation questions, data collection methods, analytic techniques, and findings. The process used by SEAs to develop capacity for gaining access to information resources and for linking such resources to the needs of educators are described.

- Intelligence for Dissemination Service Capacity: A Conceptual Framework (March 1979) expands an earlier framework into a heuristic device for studying users of educational dissemination services. This conceptual framework was completed to guide the development and refinement of questions, variables, and instrumentation for users and usages of dissemination services.

- Information About Users and Usages: A Literature Review (March 1979) is a companion document to A Conceptual Framework. The review was prepared as part of the design process used to develop the framework.

- The Client Assessment Package (December 1978) is a set of five machine-readable instruments developed by NTS to record the process of seeking and using information and assistance for educational improvement. Linked by a common identification field, the five forms in the package are the Service Form, Process Form, Linker Form, Immediate Feedback Form, and Client Assessment Form. An accompanying Guide to the Client Assessment Package provides instructions for completing and using the forms.

- Request for OMB Clearance with Supporting Documents for the Evaluation of the State Capacity Building Program in Dissemination (June 1978) is the justification and instrumentation package prepared for and approved by the Office of Management and Budget for use in the evaluation.

- A Framework for the Evaluation of the State Capacity Building Program (May 1978) presents the organizing framework for the evaluation.

During Phase I of the NTS study, five major documents were also produced:

- 1977 State Abstracts (September 1977) contains summaries of twenty-four capacity building projects. Included are the ten states initially funded in 1975 and the fourteen additional projects funded in 1976. The abstracts document dissemination activities in the SEAs as of May 1977.

- State Reports (July 1977) contains extensive documentation on nine of the first states funded through the capacity building component of the State Dissemination Grants Program. The mini-case studies examine dissemination activities in nine SEAs as of May 1977.

- A Compendium of Evaluation and Documentation Forms Currently in Use by State Capacity Building Projects (July 1977) is a compilation of selected instrumentation used by the capacity building projects. An accompanying narrative describes the included materials.

- Final Design Report for the Evaluation of the State Capacity Building Grants Program (July 1977) is a two-volume report. Volume I contains the proposed designs for the full-scale evaluation. Volume II contains proposed instrumentation.

- Building Capacity in Dissemination: Literature Review (March 1977) was used to inform the design process. The Literature Review consists of two separate but related products. The first summarizes dissemination literature, using an organizing framework which parallels that followed in NTS design work. The second product consists of an extensive bibliography and outline of topics covered in the Review. Each entry in the outline is followed by a list of relevant citations.

By describing and evaluating the process of developing dissemination capacity in selected SEAs and by assessing the program's effects, the NTS study has provided basic information for the improvement of state dissemination efforts, developed mechanisms for the continual evaluation and measurement of dissemination capacity, and by so doing, attempted to enhance the development of a nationwide dissemination system or configuration for improving educational practice and enhancing educational equity.

ACKNOWLEDGMENTS

This volume, A Study of the Development of Scales Measuring Dissemination Capacity, was prepared to supplement the evaluation of the State Capacity Building Program conducted by NTS Research Corporation for the National Institute of Education.

In Volume I, Final Report, we acknowledged those who contributed to our general efforts over the past four years; these acknowledgments highlight those who made contributions specific to the conceptualization, data collection, analysis and production of this substudy.

We would like to thank the 25 project directors and all the state personnel who participated in the study for without their assistance this study could not have been conducted.

We would also like to thank John Egermeier, Project Officer, of NIE's Research and Educational Practice Program for his valuable suggestions and critiques. A. Jackson Stenner of NTS offered valuable insights in our effort to develop the scales.

Other NTS staff also contributed to this report. Norm Frieberg assisted with the computer analyses and Dick Merritt maintained the data base. Our project secretary, Celeste Burnett together with Barbara Priboth and Lynn Elliott, with dedication and perserverance, typed and produced the entire report, remaining patient throughout numerous revisions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The purpose of this substudy is to provide documentation regarding the development and interpretation of the scales which were designed to measure the various facets of state dissemination systems. NTS Research Corporation, under Contract No. 400-76-0166, Building Capacity for Improvement of Educational Practice: An Evaluation of NIE's State Dissemination Grants Program, collected information on the activities of states as they developed dissemination capacity. These data were subjected to a scaling analysis for for the following purposes:

- To confirm that indicants grouped together on the basis of their content were, in fact, measuring the same thing.
- To eliminate indicants not related to the other indicants in the facet.
- To parsimoniously summarize a large body of data
- To serve as a measuring device, by assigning facet scores to each scale for use in characterizing state capacity.

The scales were first presented in the Interim Report of this contract (NTS, 1979). As we examined and interpreted the scales for that report, we noted a possible pattern of development of capacity in many of the scales which would indicate that a similar process was occuring across states. If so, that developmental process would be of interest to investigators of organizational development. It would be potentially useful to NIE as a tool for monitoring, for providing technical assistance, and to assist in making awards. We were unable to investigate these potential uses of the scales for the Interim Report. This substudy presents the results of NTS' further

refinement of the scales and assesses the potential utilization of the scales. The rest of this report is organized into four major sections:

Section 2, Conceptual Framework, in which the hypothesized relationships between various domains that may influence the building and utilization of capacity are described, along with the placement of the scales in this framework;

Section 3, Scaling Methodology, where the procedures used to develop the scales, including the selection of items for the scales and the assignment of scores for each state on the scales, are discussed; followed by an interpretation of each resultant scale.

Section 4, Capacity Building as a Developmental Process, where the degree to which the scales provide a developmental perspective to the capacity building process is assessed; and

Section 5, Interpretation and Utilization, where the interpretations that can be drawn as well as those which should not be drawn from the scales are discussed. How the interpretations might be utilized for purposes of program improvement is also discussed.

The first half of Section 3, Development of the Scales, is a relatively technical discussion of how the scales were developed and how Rasch scaling analysis was employed. For those readers less interested in this technical discussion we suggest that this half could be omitted without losing the essence of the substantive discussion. We do suggest, however, that the second half of the section, Interpretation of the Scales, be reviewed for an understanding of the content of the scales.

There are three companion volumes to this report:

Building Capacity for Educational Improvement. An Evaluation of NIE State Dissemination Grants Program.

1979 State Abstracts, a document which profiles dissemination activities in thirty-eight SEAs; and

A Study of Linker Agent Activities and Roles.

CONCEPTUAL FRAMEWORK

As a part of the approach to meet the basic objectives of this study, a conceptual framework was developed which specified and organized the variables which define the process of building dissemination capacity. The conceptual framework is presented in Figure 2.1. The framework includes five components and specifies the relationships between the components. The five components of the framework are: (1) State and SEA Contextual Characteristics; (2) NIE Program Characteristics; (3) State Capacity Building Project Characteristics; (4) Facets of an SEA Dissemination System; and (5) Outcomes: System Outputs and Impacts. These components can be further classified into three major categories: Context, Process, and Outcomes.

The Context Domain

Component 1: Contextual Characteristics. The first component refers to variables which describe contextual characteristics of the particular state and SEA. State characteristics include such variables as state size, existence and use of intermediate service agencies or regions, school enrollment, number of school districts, and modernity (Herriott and Hodgkins, 1973) SEA characteristics include such variables as attitudes in the SEA towards change, centralization (Wirt, 1977), previous and current involvement of the SEA in dissemination activities, and the relative influence of the SEA, intermediate education agencies (IEAs), and LEAs in local educational improvement.

Component 2: NIE Program Characteristics. The second component presents factors which characterize the NIE Program, including its design and operations at the Federal level. Included are Program goals, Program

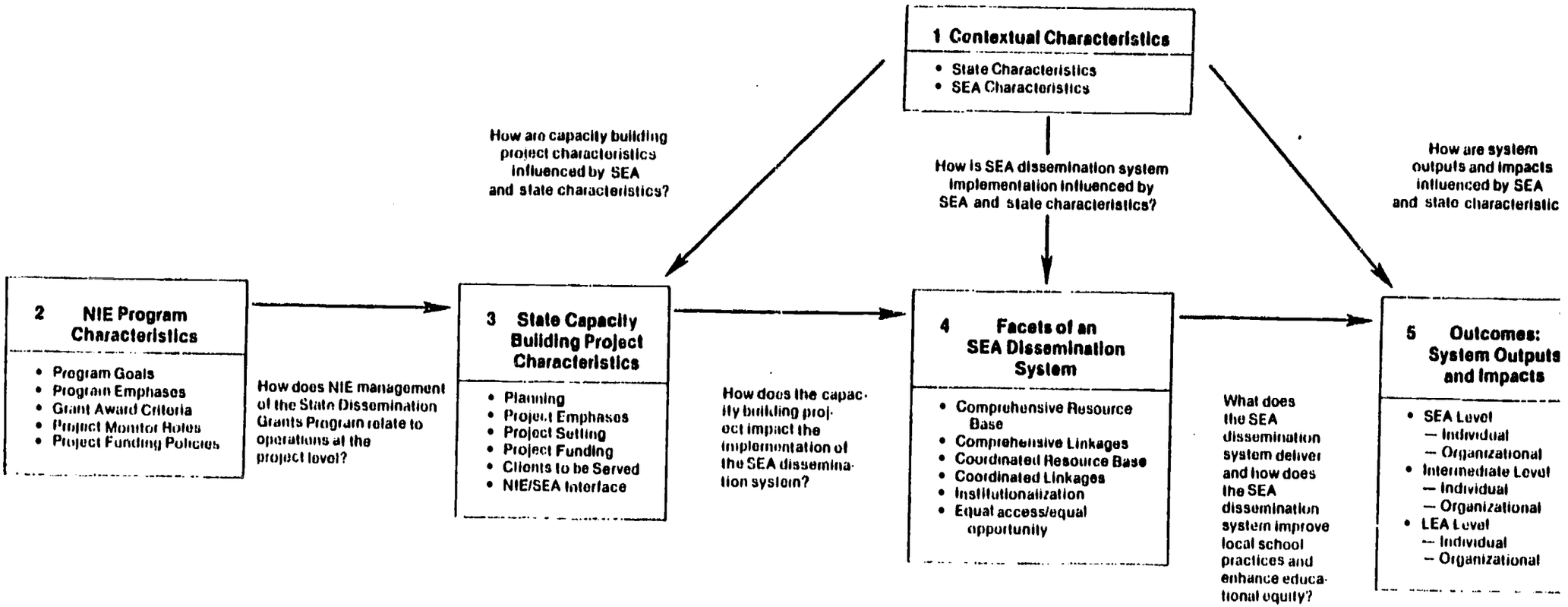


FIGURE 2.1 A Framework for the Evaluation of the State Capacity Building Program

emphases, grant award criteria, project funding policies, and project monitor roles.

The variables contained within Components 1 and 2 of the conceptual framework may be considered as the context domain, which includes both the legal/policy framework and the social/political setting within which the projects are situated. The legal/policy framework includes not only the SCBP and guidelines associated with the Program, but also other Federal and state dissemination programs and policies.

The Process Domain

Component 3: State Capacity Building Project Characteristics. The third component contains variables which describe the resultant program intervention at the SEA level; that is, the project's structure and activities. Project structure includes such variables as the project's number of years in the SCBP, funding, location within the SEA, project director tenure, and management arrangements. Project activities include such factors as targeting clients for dissemination activities, building additional information files, working with IEAs to provide linkage to local school districts, and interacting with NIE.

Project characteristics may be considered as comprising the process domain, and include project structures and activities as well as other SEA dissemination activities. It must be noted that the actions taken by the state and the SEA are activities which tend to obscure the relationship between the project and the building of dissemination capacity. Phrased in another way, the project is provided "leverage" money through which a wide range of activities are generated, activities which often extend beyond the project. At the same time, the state/SEA may also be developing additional activities outside of the project which enhance the dissemination system.

The Outcomes Domain

The outcomes domain includes two categories of variables within the conceptual framework: 1) Facets of an SEA Dissemination System; and 2) Dissemination System Outputs and Impacts. The outcomes domain includes not only those variables which are appropriate objectives for this study to measure, that is growth of dissemination capacity, but also those which reflect the long range goals of Federal policy as stated by NIE.

Component 4: Facets of an SEA Dissemination System. This component refers to those elements which comprise a dissemination system: the resources, linkers, linkages, services, and institutional arrangements an SEA develops, implements, and institutionalizes to improve local educational practice and enhance educational equity. An SEA dissemination system, of which the capacity building project is usually just a portion, is depicted as being comprised of six facets. Facet 1, Comprehensive Resource Base, refers to the types of resources (e.g., ERIC, promising practices files) that the SEA has the ability to access. Facet 2, Comprehensive Linkages, refers to the availability and use of a variety of individuals (e.g., SEA staff, IEA staff, LEA staff) and media (e.g., radio, television, publications) to connect educators with the information and services needed to improve local school practice. Facet 3, Coordinated Resource Base, refers to the extent to which various mechanisms for coordinating resources (e.g., a formal referral process, knowledge by resource base personnel of other resources, use of other resource bases) are available and used. Facet 4, Coordinated Linkage, refers to the extent to which linkers coordinate their activities with resources and linkages in order to provide assistance to improve local school practice.

Facet 5, Institutionalization, refers to the extent to which the SEA dissemination system is developed so that it remains after NIE funding of the SEA's capacity building project terminates; the provision of funds by the state for dissemination and statements by the chief state school officer supporting dissemination are examples of variables included in Institutionalization. Facet 6, Equal Access/Equal Opportunity, refers to the extent to which the SEA dissemination system provides resources to all educators on all topics, and the targeting of dissemination efforts to individuals who work with special populations, including the handicapped, minorities, and women.

Component 5: Dissemination System Outputs and Impacts, refers to outputs of the SEA's dissemination system and their effects on the improvement of local practice and enhancement of educational equity. These systems outputs and impacts relate to the longer range goals of improving local educational practice and enhancing educational equity. An investigation of these outcomes was beyond the scope of the overall study; however, when possible, we presented descriptions in Volume 1, Final Report, indicative of the effect of dissemination capacity building upon these outcomes.

Uses of the Scales

Having described the relationship of the dissemination system scales to the conceptual framework, we suggest that the scales can serve an important purpose for this evaluation of the Program. The scales provide us with measures of a dissemination system. Through the use of the scales we can conduct analyses which search for those factors that influence the extent to which the dissemination capacity is developed and the dissemination system implemented. We will describe in the following chapters the development of the scales and investigate their ability to assist us in assessing dissemination capacity and implementation.

SCALING METHODOLOGY

Development of the Scales

This chapter describes the procedures used to develop a scale for each facet of the SEA dissemination system and to assign scale scores for each state. The purpose of each scale was to allow measurement of all of the states on a particular facet. However, for this to occur, the scores of the states on each scale had to be comparable to one another. The ability to compare states was important, as one of the major purposes of the study was to assess what impact state and project characteristics had on making one state different from another with respect to the building of dissemination capacity. Thus, the indicants (i.e., individual items) that were selected for each facet had to be measuring the same thing (e.g., comprehensiveness of resource base) for all of the states. This meant that indicants that did not appear to be measuring the same thing had to be eliminated, even though they might reflect an important part of one or two states' dissemination systems. It was necessary to measure what the states had in common with respect to each facet, and to ignore, therefore, what might be a unique approach that some states were taking towards building capacity. The elimination of such indicants in no way represented a judgment as to the appropriateness of a state's approach. Rather, it simply meant that those indicants were not reflective of the facet that the rest of the indicants in the scale were measuring. It would have been inappropriate to combine an indicant that was measuring one thing with a group of indicants that were measuring something else in the computation of a facet score.

Each scale was comprised of the set of indicants that best measured that facet. The indicants and states were then ordered on the scale. A state's

placement on a scale reflected the degree to which a facet had been developed (i.e., the attainment of resources or activities) by that state.

Procedures

In order to produce a quantitative description of each state's dissemination system with respect to these six facets, it was necessary to develop scales that would measure a state's position on each of the scales. These scales were based on the responses of SCBP directors to the Capacity Building Indicants questionnaire (CBI) that was administered in 1978 and 1979. The responses from the 1979 administrative were used to verify the analyses based on the 1978 data. Since all of the directors indicated that no restrictions existed with respect to Equal Access/Equal Opportunity, there was no need to develop a scale for this facet; all of the states would have fallen at the highest point on this scale.

The scale for each facet was developed using a three-step process, with the second and third steps being repeated until a final set of indicants was identified that measured that facet and only that facet. In the first step, indicants were selected and grouped into scales on the basis of their content. In the second step, the empirical interindicant relationships were used to eliminate those indicants in each scale that did not seem to be measuring the same dimension as the other indicants in the scale. In the third step, the scores of the states on the individual indicants were submitted to a latent trait scaling analysis, and if necessary, additional indicants were eliminated.

Step One: Content analysis. In the first step, a content analysis of the available indicants was conducted to determine the set of indicants which appeared to measure the same facet. The criteria for inclusion in this step

were fairly liberal, in order not to eliminate an indicant that might be strongly related to other indicants even when this was not immediately obvious from the indicant's content.

Step Two: Interindicant relational analysis. In the second step, the interindicant correlations were used in a reliability analysis to assess the degree to which the indicants in each scale were all measuring the same facet. The coefficient alpha (KR_{20}) reliability was computed for each scale. Two measures were employed to assess the extent to which a particular indicant was measuring the same facet as the other indicants. First, correlations were computed between each indicant and the total of all the other indicants in the scale. Second, KR_{20} coefficients were computed excluding each indicant, in turn. Indicants which had low indicant-total correlations or which lowered the reliability of the scale were excluded.

Step Three: Scaling analysis. The indicants retained after the reliability analysis were submitted to a latent trait analysis using the Rasch (1960) scaling model. This model postulates that the probability of a state having or not having a particular indicant (i.e., a positive response or a negative response) is a function of the difference between the state's score on the scale and the indicant's score on the scale. If the state's score is higher, the probability of a positive response will be greater than .5 (i.e., greater than a 50/50 chance), while if the state's score is lower, the probability of a positive response will be less than .5 (i.e., less than a 50/50 chance). The state's score on a scale reflects how much coordination, comprehensiveness, or institutionalization its dissemination system has, and is based on the state's total number of positive responses with respect to the indicants of the scale. The indicant's score on the scale reflects how

difficult it is for any state to have a positive response on that indicant, and is based on the states' total number of negative responses with respect to that indicant. For example, the scale measuring the comprehensiveness of the resource base had a number of resources arranged on the scale, with the ordering of the resources directly related to how many states possessed each of those resources.

For a state to be relatively high on the scale, therefore, it must have positive responses to more of the indicants than most of the other states. This is an indication that the state is higher on the facet than are most of the other states. For the indicants, this relationship is reversed. The more states that have negative responses on a particular indicant, the higher the scale value of that indicant will be. For an indicant to have a large number of positive responses implies that the indicant cannot be too high on the facet being measured because so many states possess that indicant. Indicants with very few positive responses, on the other hand, will be highest on the facet, reflecting the fact that very few of the states possess a sufficiently high amount of the attribute in order to score positively on the indicant.

The Rasch model analyzes all of the states' responses to all of the indicants and places the indicants and the states on the same scale. Since the Rasch analyses are based on positive/negative responses, all of the indicants were dichotomized, except for the Comprehensive Resource Base indicants, which were already dichotomized as provide/do not provide. The responses to the other indicants were based on a five point scale, with a "1" meaning "very limited," or "nonexistant" or "never," and a "5" meaning "very extensive" or "always." To dichotomize these indicants, we combined the "very

limited" and "limited" categories into negative responses, and the "moderate," "extensive," and "very extensive" categories into positive responses.

The dichotomized responses were then submitted to a Rasch scaling computer program called BICAL (Wright and Mead, 1979). In addition to computing scale values for both the states and the indicants, this program also provides several measures of the degree to which the indicants are all measuring the same facet. The first of these statistics, the "total t-test," measures the extent to which the predicted responses for that indicant matches the observed responses for that indicant. The higher the value, the lower the match. Wright and Mead suggest a value of 2.0 as the cutoff for this total t-test statistic, beyond which, they say, the statistic is indicating a poor fit of the predicted responses for that indicant to the data. A second statistic, the "between t-test," is based on dividing the states into two groups, those with high scores and those with low scores, and assessing the degree to which the predicted average response to an indicant for each group matches the observed average response to the indicant for each group. As with the total t-test statistic, the higher the value, the poorer the fit; again we followed Wright and Mead's suggestion that a value of 2.0 be used as a cutoff. All indicants with either a total t-test statistic or a between t-test statistic of greater than 2.0 were eliminated.

Steps Two and Three were repeated for the remaining indicants. This was necessary because whenever some of the indicants were eliminated, the statistics for the remaining items were altered. Since the statistics all reflect the degree to which an indicant is measuring the same facet as the other indicants, they are sensitive to what indicants are placed in the "other" category. Consequently, an indicant might not have originally detracted from the

scale's reliability or the scale's fit to the Rasch model within the complete set of indicants. However, that same indicant might detract from the reliability or the fit of the scale when that scale is based upon a subset of the original indicants. If all of the indicants again passed all of the various tests of fit, then the process stopped. If not, then the appropriate indicants were eliminated, and Steps Two and Three were repeated once again. For most of the scales, only two iterations were required, and in no cases were more than three iterations necessary.

When the final Rasch scaling analysis was completed, we examined the relationship between the raw scores of the states on each scale and the state scale values as determined by the Rasch analysis. We noted that the relationship between the total "raw" scores and the scale scores was almost linear; i.e., a change of 1 raw score point for each scale tended to correspond with a change of .2 on the Rasch scale for virtually the entire range of raw scores. Consequently, for purposes of interpretability, we decided to use the actual raw scores rather than Rasch scale scores in subsequent analyses.

The output of the Rasch scaling analysis is an ordering of the states and the indicants on the same scale. The ordering of the states is fairly straightforward to interpret: A state is higher or lower than another state, which in turn implies that the state has more or less dissemination capacity (at least with respect to that facet) than does the other state. We must emphasize, however, that this does not in any way place a value judgment on the position of the state on the scale, a point that will be discussed at length in Section 4.

With respect to the ordering of the indicants, the scale may suggest some sort of developmental paradigm. It must be emphasized, however, that

the existence of a developmental paradigm is only suggested by the scale, and not confirmed by it. What the scale tells us is that if a state has a positive response with respect to a particular indicant, then it is likely to have positive responses with respect to all of the other indicants having lower scale values. It may well be that this means that for a state to have a certain resource on the Comprehensive Resource Base scale, for example, it was necessary for it to first acquire other resources preceding it on the scale. On the other hand, the scale may simply be hierarchical, reflecting the possibility that some resources are easier to acquire than others, and that the number of resources that a state has may simply be a matter of the size of its budget rather than its position in some sort of developmental process. An empirical assessment of the developmental versus hierarchical paradigm was undertaken, and is described in Section 4.

The 1978 and 1979 indicant responses for the various scales are displayed in Figures 3.1 to 3.6. The institutionalization scale is displayed in Figures 3.6a and 3.6b. The scale is unitary (i.e., it is one scale) but it is divided in order to express important information about the institutionalization process. The indicants in Figure 3.6a are related to what we believe occurs during the initial stages of capacity building and centers upon project activities, while the indicants in Figure 3.6b are related to agencywide activities that occur later in capacity building. The figures display how many of the 25 states responded positively to each indicant in 1978 and in 1979. The total t-test fit statistics for the indicants are shown next to the indicant descriptions. For the 1978 scales, all the t-test values for the indicants fall well below the 2.0 cutoff criterion suggested by Wright and Mead; the same is true for virtually all of the indicants in the 1979

scales as well. For those indicants to which either all of the states or none of the states had positive responses, the Rasch model will fit perfectly and therefore the fit statistic is meaningless. For these indicants, the fit statistic has been replaced with "**".

The reliabilities for each of the six scales were calculated for the 1978 and 1979 data. The KR_{20} internal consistency formula was used, and the reliabilities are presented at the bottom of Figures 3.1 to 3.6. While it is true that reliability values can range from 0 (no internal consistency) to 1 (perfect internal consistency), the interpretation of reliability coefficients must be tempered by an awareness of the number of indicants in the scale. The more indicant that a scale has, the more reliable it will tend to be. The nature of the relationship between reliability and test length is stated in the Spearman-Brown prophecy formula:

$$r' = \frac{mr}{1 + (m-1)r}$$

where r is the reliability of the original test, m is the multiple of the original test length by which the test has been lengthened or shortened, and r' is the reliability of a test m times as long as the original. This formula can be used to re-express the observed reliabilities of the original scales in terms of scales that have a certain uniform number of indicants. In order to provide a context for comparison, we chose 30 as the number of indicants, since this is a typical number of items that appear in a nationally standardized test or subtest (e.g., reading vocabulary, reading comprehension, mathematics computation, etc.). Tests or subtests of such length are generally considered to be sufficiently reliable if their KR_{20} internal consistency measure equals approximately .90.[3.1]

Therefore, we have

3.1 See, for example, Technical Bulletin 1 for the 1977 edition of the California Achievement Test, where a substantial proportion of the subtests have about 30 indicants and a KR_{20} reliabilities of about .90.

presented in Figures 3.1 to 3.6 the reliability for a 30 indicant scale based upon the observed reliabilities of each scale in 1978 and 1979. These 30 indicant reliabilities all tended to be quite high, indicating that the internal consistency of these scales was on a par with the internal consistency of commercially available, nationally standardized tests and subtests. It might be noted that it would have been possible to select the indicants in the scales such that these reliabilities would have been even higher. However, our primary criterion for indicant selection was that the indicants total t-fit fall below 2.0. While these two approaches will tend to yield similar sets of indicants, they will not be in perfect agreement. The reliabilities displayed in Figures 3.1 to 3.6, however, show that even though a Rasch scaling criterion was used for purposes of indicant selection, the scales also did quite well with respect to the reliability measure associated with more traditional test theory procedures. Thus, it is clear that for each scale, the selected indicants all tended to be measuring the same thing.

We examined the relative positions of the 1978 indicants to the relative positions of the 1979 indicants. The absolute positions of the 1979 indicants tended to be somewhat higher, reflecting that the 25 Cohort I, II, and III states had been in the program one year longer. However, we still expected the relative positions to be the same. Consequently, we computed a Spearman rank order correlation coefficient for each scale comparing the 1978 indicant ranks (i.e., the relative position of an indicant with respect to the other indicants in the scale) with the 1979 indicant ranks. These correlations are displayed at the bottom of each figure. These correlations are all quite high, indicating that the ordering of the indicants along each scale had not changed appreciably from 1978 to 1979. The fact that the

ordering remained reasonably invariant is extremely important, as it implies that the set of indicants constituting each scale were measuring the same thing at both points in time. Not only does this mean that the indicant structure was replicated, but that it is meaningful to talk about the differences in the 1978 and 1979 scores as reflecting change, since the same facet was being measured in the same manner both times.

Interpretation of the Scales

Having now discussed the procedures used to develop the scales and the psychometric properties of the scales, we now consider what sort of interpretations can be attached to the scales. Such an interpretation must necessarily be based on an examination of the selected indicants and their positions on the scales, but it is also relevant to consider what indicants were not selected as well. Accordingly, we present all of the indicants associated with each scale in Tables 3.1 through 3.6. In each table, the indicants associated with the primary scale are shown first, followed by secondary scales that include groupings of unselected indicants which tended to be associated with one another, as evidenced by their interindicant correlations. These are followed by a list of the remaining indicants that appeared to have little relationship with any of the other indicants.

It should also be noted that our discussion reflects an assessment of the content of the scales as it relates to the developmental paradigm; as we discuss the ordering of the indicants on each scale, we are also discussing the ways in which the states might proceed in the development of their dissemination systems.

Comprehensive Resource Base

The breakdown of the Comprehensive Resource Base scale into the various indicant subgroups is shown in Table 3.1. The primary comprehensive resource

TABLE 3.1

COMPREHENSIVE RESOURCE BASE

PRIMARY SCALE

1. ERIC - Education Resources Information Center
2. ECER - Exceptional Child Education Resources
3. NICSEM/NIMIS - National Information Data Organizer
4. NTIS - National Technical Information Service
5. Federal legislation file
6. Promising practices file
7. Listings/descriptions of federal and state funded innovative programs
8. Listings/descriptions of operating local exemplary initiatives
9. SEA human resource file
10. IEA human resource file
11. LEA human resource file
12. State legislation file
13. SEA products
14. NDN products
15. PREP Packages - Putting Research into Educational Practice

SECONDARY SCALE #1

1. AIR/ARM - Abstract of Instructional and Research Material in Vocational and Technical Education
2. Dissertation Abstracts
3. Lab and center products

SECONDARY SCALE #2

1. Education Index
2. FIDO - Fugitive Information Data Organizer
3. Files on user needs in the state

SECONDARY SCALE #3

1. PA - Psychological Abstracts
2. Sociological Abstracts
3. SSCI - Social Science Citation Index (also called Social Scisearch)

SECONDARY SCALE #4

1. State-produced instructional materials
2. Institutes of higher education-produced instructional materials
3. Locally-produced instructional materials
4. IEA-produced instructional materials

REMAINING INDICANTS

1. PIPS - Project Information Packages
2. Right-to-Read Packages
3. NYT/IB - New York Times Information Bank

Possible responses were: Available, Not Available

base scale includes the various knowledge resources within the resource base that a project accesses in responding to clients' requests. These include:

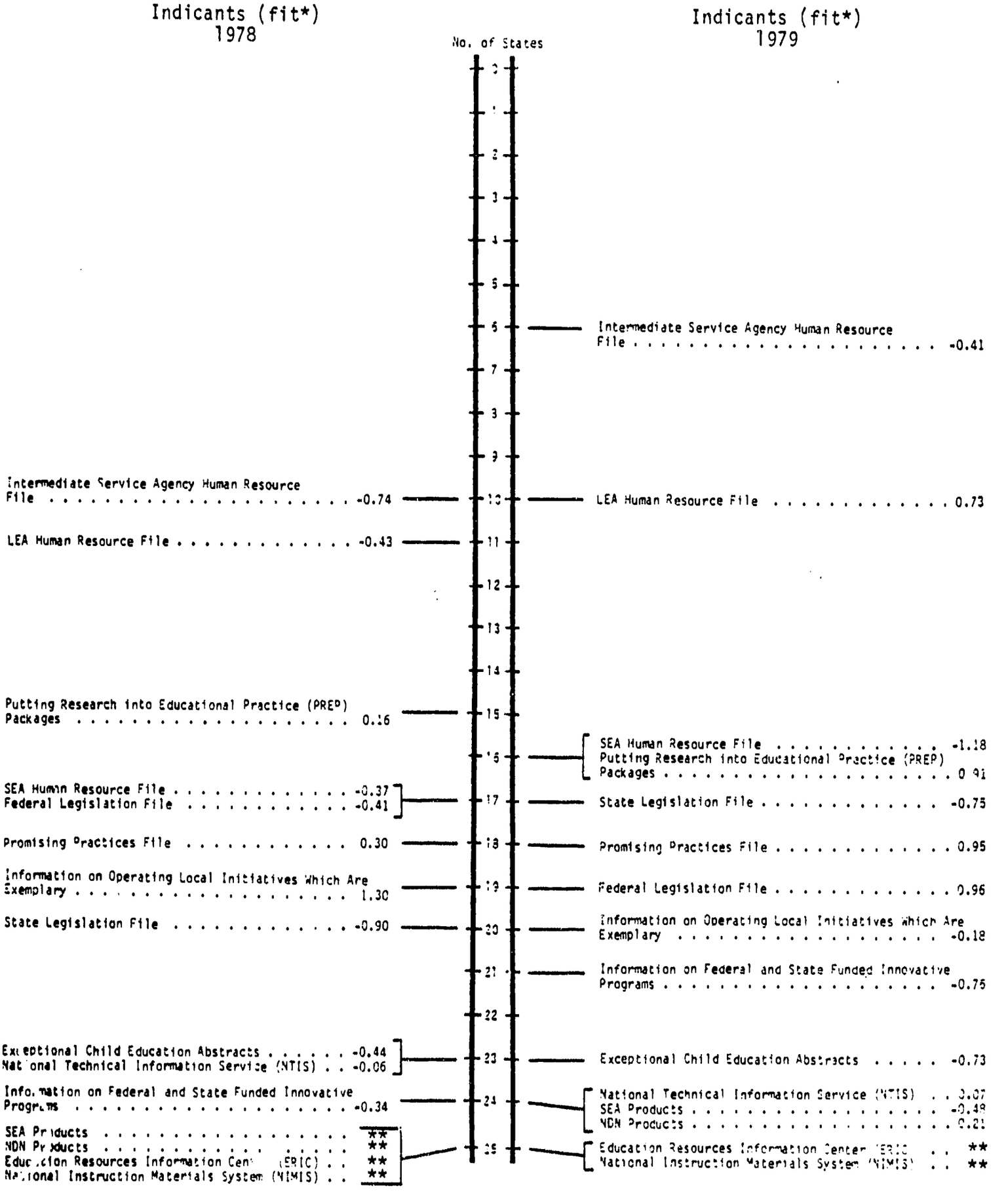
1. ERIC and other national information files;
2. Validated programs, including Federal and state-funded innovative programs and local exemplary programs;
3. Promising practices/program files, which refer to non-validated educational methods and ideas;
4. Human resource files which match client need with available consultants in the SEA, IEA, LEA, or IHE, and
5. Federal and state legislative files.

Examination of Figure 3.1 shows that in 1978 all projects had access to four resources: Education Resource Information Center (ERIC), National Information Center for Special Education Materials/National Instructional Materials System (NICSEM/NIMIS), National Diffusion Network (NDN) products, and SEA products. With the exception of NICSEM/NIMIS, the other three are resources that SEAs generally have available before funding. These four resources form a base upon which a comprehensive set of resources is further developed. Other resources typically available are descriptions of Federal and state-funded innovative programs and such files as National Technical Information Service (NTIS) and Exceptional Child Education Resources (ECER). Resources that about fifteen to twenty of the projects have are: files of promising practices, including local exemplary programs and Putting Research into Educational Practice (PREP) packages; legislative files; and SEA human resources. The least frequently included resources (utilized by less than half of the states) are LEA- and IEA-generated local human resource files.

The ordering of the indicants included in the comprehensive resource base shows two underlying patterns: One pattern involves the type of resources; the second pattern involves the sources of the materials. The most

Figure 3.1

COMPREHENSIVE RESOURCE BASE



Scale Reliability = .72
Scale Reliability for 30 Items = .84

Scale Reliability = .52
Scale Reliability for 30 Items = .68

Spearman Rank Correlation = .90

(Total t-test statistic - should be less than 2.0)
(Perfect fit)

The most widely used resources are: national data files, print-based materials, and products. Innovative programs (including Federal and state-funded innovative programs and local exemplary programs) are the next numerous. Promising practices files and legislative files are included with less frequency, while human resource files are the least often included resource in an SEA's comprehensive resource base. With respect to source, national resources are most often a part of the dissemination system, followed by state resources, local district resources, and finally intermediate education agency (IEA) resources. Notable for their absence are resources in institutions of higher education (IHEs).

The first secondary comprehensive resource base scale includes AIR/ARM, lab and center products, and Dissertation Abstracts. The second secondary comprehensive resource base scale includes FIDO (Fugitive Information Data Organizer), lab products, and files of user needs. This configuration of resources tends to occur in systems that have a relatively small number of resources, suggesting that several states acquire this cluster of resources in conjunction with ERIC to form an alternative kind of resource base, rather than acquiring the kinds of resources found on the primary scale.

The third secondary comprehensive resource base scale includes Psychological Abstracts, Sociological Abstracts, and Social Sciences Citation Index. All of these are bound abstracts and therefore do not represent resources that are distributed to clients but are instead utilized by resource base and project staff to locate relevant materials.

The fourth secondary comprehensive resource base scale includes instructional materials from various sources within the state: SEA, IEA, LEA, and IHE.

The remaining indicants include federally-prepared packages of which some are now obsolete, such as Project Information Package (PIP).

Coordinated Resource Base

The breakdown of the Coordinated Resource Base scale into the various indicant subgroups is shown in Table 3.2. The primary coordinated resource base scale describes a spectrum of behavior ranging from a broadening of awareness on the part of project staff of the various components of the resource base in the SEA and LEAs to the development of referral and contact procedures to assure the availability of all extant resources for meeting client requests. Examination of Figure 3.2 shows that for both years, almost all of the states reported that the resource base and project staffs were aware of the components of the resource base, so that these appear to represent the baseline for describing a coordinated resource base. As one reads up the scale, from those indicants which are used by more states to those used by fewer states, and if one assumes that those elements used in more states are easier to develop than those used by fewer states, one can envision the process of coordinating resources. Thus the following scenario, based on the ordering of the indicants and supported by our site visits, is suggested.

At the earlier stages of coordination, the services to clients tend to overlap and the various resource bases duplicate each other's efforts. Operationally, a client could access a variety of sources for the same information. This does not necessarily mean, however, that the centralized resource base had as comprehensive an information file in the area of special education, for instance, as the special education unit had.

As the SEA and the other resource agencies become aware of the services the project can provide and cooperation between the project and other

TABLE 3.2

COORDINATED RESOURCE BASE

PRIMARY SCALE

1. Resource base staff awareness of components of comprehensive resource base
2. Project staff awareness of components of a comprehensive resource base is
3. A formal referral process that incorporates procedures to avoid duplication of effort
4. Responses coupled with one or more referrals are
5. The rate of rejections to responses are
6 The number of contacts/referrals of requests to other source agencies are
7. Usage of the compendium (documented compilations) of resources by other resource agencies is

REMAINING INDICANTS

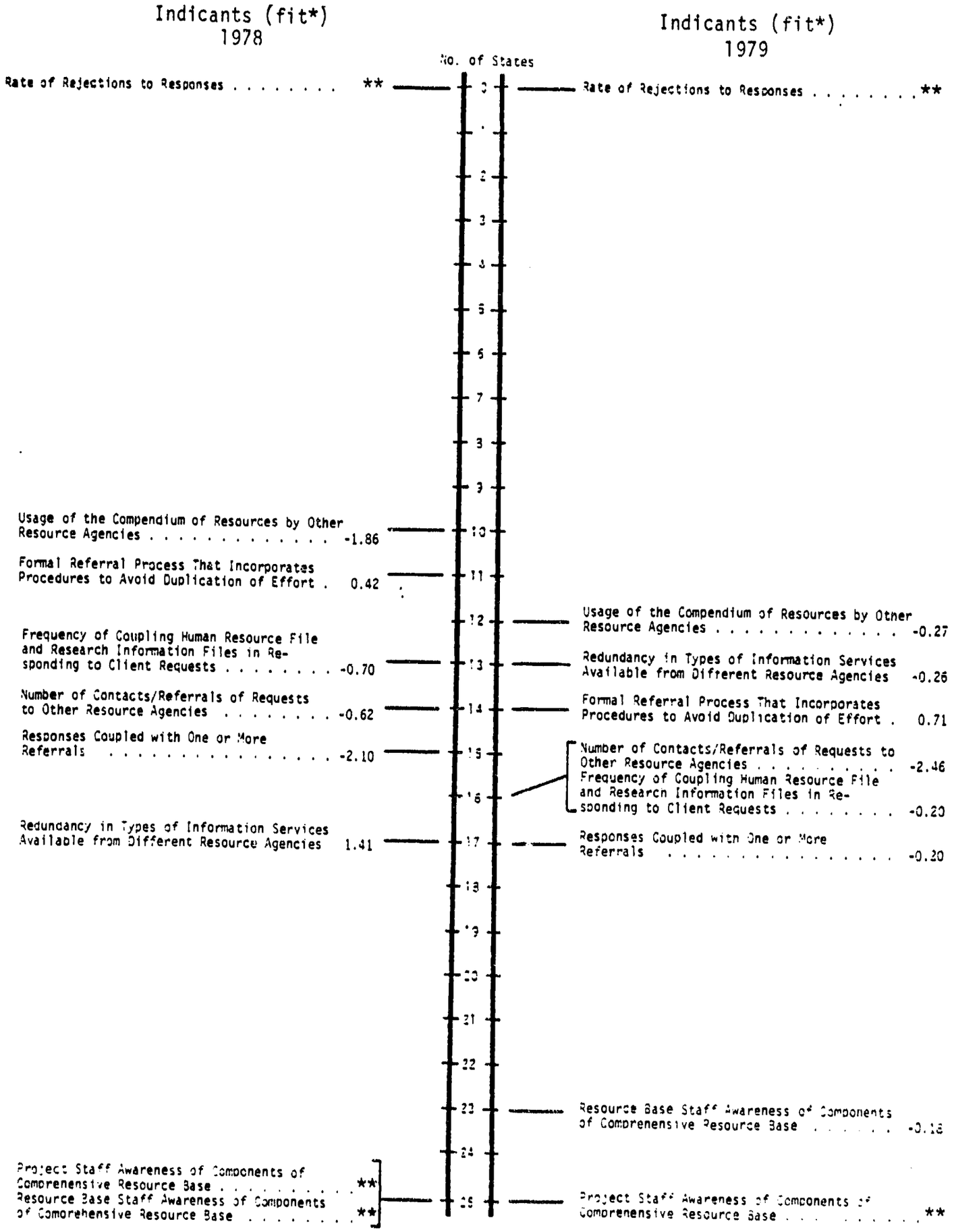
1. SEA inner circle of administrators awareness of a comprehensive resource base is
2. Other resource agencies (IHE's, museums, state system) awareness of project staff and services is
3. Resource base service to Title IVC, special education, vocational education is
4. Other resource agencies (IHEs, museums, state library system) usage of project staff or services is
5. Project awareness of other sources contacted by client is
6. Redundancy in types of information services available from different resource agencies is
7. Communications between linkage agents and resource base(s) are
8. Multiple source response to single source (e.g., ERIC) responses are
9. The frequency of coupling a human resource file and research information files in responding to client requests is
10. Agencywide understanding of a common definition of dissemination is

---

Possible responses were: Non Existent, Limited, Moderate, Extensive, Very Extensive

Figure 3.2

COORDINATED RESOURCE BASE



Scale Reliability = .61
Scale Reliability for 30 Items = .85

Scale Reliability = .46
Scale Reliability for 30 Items = .76

Spearman Rank Correlation = .75

(*Total t-test statistic - should be less than 2.0)
(**Perfect fit)

resources is achieved, the central resource base expands its services and broadens the variety of materials which can be included in the response to a client's request. This is accomplished by coupling responses with referrals to other agencies which may have more extensive sources of information pertaining to the client's request. The development of coordination/cooperation is further achieved when the central resource base begins to contact these other agencies for the client rather than referring the client to the other agency. At this point the response sent from the project to the client includes information from a variety of sources.

This "reaching out" of the project to other agencies or program units develops working relationships which form for the SEA a "compendium of resources," a network of units which begin to share information. This sharing, as shown in our site visits, can take the form of cooperative agreements between program units to respond to these requests for information or it can take the form of other program units providing the information sources to the central resource base for inclusion directly into centralized files.

As greater coordination is achieved, other resource agencies become aware of the project and its services and begin to utilize the project to collect information for their clients. Finally, a formal referral process is developed through which the SEA can respond with a minimum of duplication of effort. In practical terms, when there is adequate coordination of the SEA resource base, a client can request assistance from any program unit in the SEA network and receive a comprehensive (i.e., data from multiple sources) response.

The remaining indicants form a hodgepodge. They include indicants that were rejected before scaling procedures because they either: (1) unnecessarily duplicated items in other scales (i.e., "agencywide understanding of a

common definition of dissemination" was also included in institutionalization); (2) were vague and subject to differing interpretations (i.e., "multiple source response for single source responses"); or (3) did not pertain to all the states (i.e., "coupling a human resource file with a research information file" would not indicate coordination between resource bases if the human resource file was also included in the resource file or if the state did not have a human resource file). Other indicants probably did not correlate with the primary scale because they focused on the awareness and usage by the other resource agencies of the project's resource base, whereas the primary scale's focus is on the project's actions in coordinating and using these agencies.

Comprehensive Program Linkage

The breakdown of the Comprehensive Program Linkage scale into the various indicant subgroups is shown in Table 3.3. The primary comprehensive program linkage scale includes three groups of linkages that appeared to reflect relationships which were interorganizational in character:

1. dissemination specialists, including resource base staff, NDN staff, and Title IV staff;
2. program-specific specialists, including the staffs of such programs as: special education, career education, Title I, handicapped education and early childhood education; and
3. state library system staff.

Examination of Figure 3.3 shows that for both years, the most frequently developed linkages are resource base staff, Title IV-C, and NDN staff, who are often referred to as "dissemination-type specialists." The least developed linkage elements are with the staffs of the state library system and migrant education and early childhood education programs.

TABLE 3.3

COMPREHENSIVE PROGRAM LINKAGE ELEMENTS

PRIMARY SCALE

1. Resource base staff
2. Title I staff
3. Special education staff
4. NDN staff
5. Title IVC staff
6. Early childhood education staff
7. Career education staff
8. Migrant education staff
9. Handicapped education staff
10. State library system staff

SECONDARY SCALE #1

1. LEA representatives
2. Building level representatives
3. Local school librarians
4. School board members

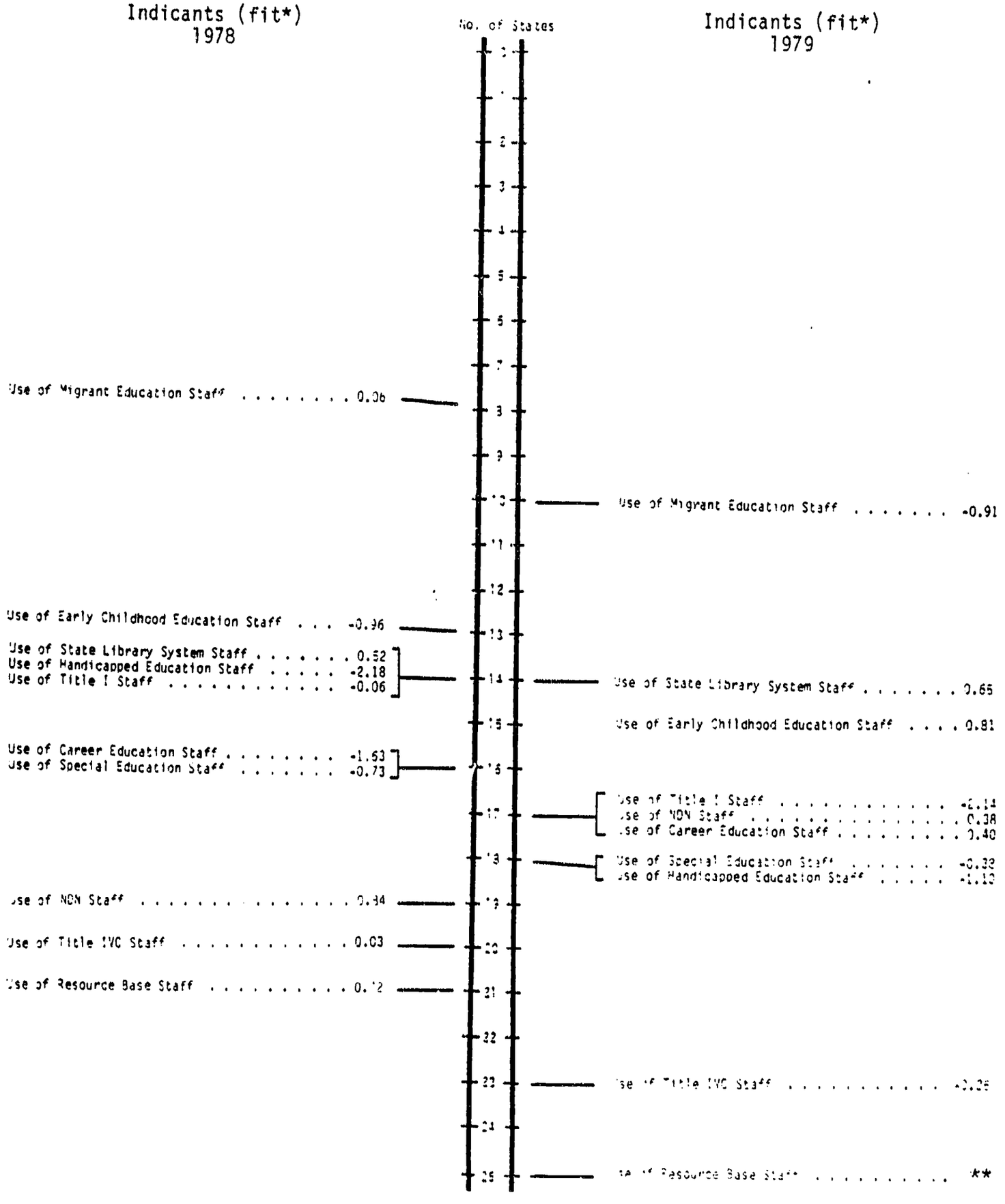
REMAINING INDICANTS

1. State associations (content-oriented) staff
2. Other state agency staff
3. Regional education center staff
4. Intermediate service agency staff
5. Vocational education staff
6. Adult education staff
7. Content-oriented specialists (SEA)
8. Urban education staff
9. Right-to-Read staff

Possible responses were: Unavailable, Not Used, Limited Use, Moderate Use, Extensive Use, Very Extensive Use

Figure 3.3

COMPREHENSIVE PROGRAM LINKAGE ELEMENTS



Scale Reliability = .79
Scale Reliability for 30 Items = .92

Scale Reliability = .70
Scale Reliability for 30 Items = .88

Spearman Rank Correlation = .81

(*Total t-test statistic - should be less than 2.0)
**Perfect fit)

Several trends can be identified between the 1978 and 1979 scales. One trend that seems clear is that projects generally first involve elements that are close to it (i.e., at the SEA) and then proceed to enlist persons at levels closer to local education, most notably intermediate education agency staff. While the involvement of NDN staff experiences a slight decline, the involvement of Title IV staff and many program-specific staff generally increases. Possibly as a result of further implementation of Public Law 94-142 (Education for Handicapped Children Act), special education and handicapped education staff showed the most dramatic increase in involvement, followed by Title I staff.

The secondary comprehensive program linkage scale includes linkages at the local levels, including LEA and building level representatives, local librarians, and school board members. While the primary scale reflects interorganizational linkages, the secondary scale represents local linkers who serve linker functions.

The remaining indicants include: (1) content-oriented specialists, including SEA staff and members of content-specific state associations; (2) regional staff; and (3) staff of such programs as vocational education, adult education, and urban education. It may be that content-oriented specialists and particular program staff are more difficult to establish relationships with due to "turf" battles. Regional staff is a poor indicant since it is redundant with IEA level staff represented in the primary scale.

Comprehensive Media Linkage

The breakdown of the Comprehensive Media Linkage scale into the various indicant subgroups is shown in Table 3.4. The primary comprehensive media linkage scale includes print-based materials and electronic devices that are

TABLE 3.4

COMPREHENSIVE MEDIA LINKAGE ELEMENTS

PRIMARY SCALE

1. Project-specific publications
2. SEA publications
3. Newspapers
4. Educational television
5. Audiovisual aids
6. Computer-based user systems
7. Slides
8. Films
9. Prerecorded cassettes

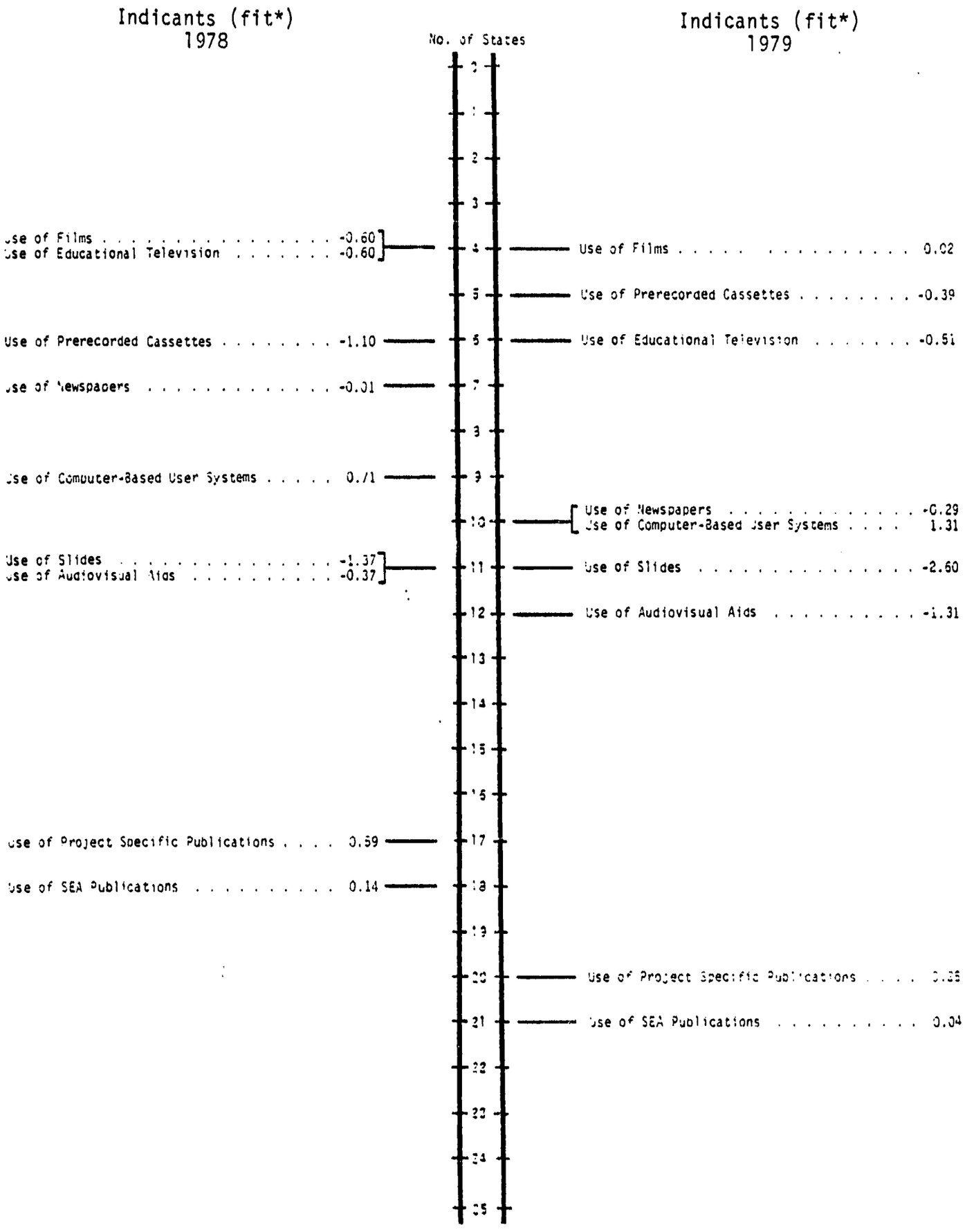
SECONDARY SCALE #1

1. Closed circuit television
2. Commercial television
3. Radio

Possible responses were: Unavailable, Not Used, Limited Use, Moderate Use, Extensive Use, Very Extensive Use

Figure 3.4

COMPREHENSIVE MEDIA LINKAGE ELEMENTS



Scale Reliability = .80
Scale Reliability for 30 Items = .93

Scale Reliability = .63
Scale Reliability for 30 Items = .84

Spearman Rank Correlation = .96

(Total t-test statistic - should be less than 2.0)
Perfect fit)

utilized by projects to create awareness and interest in the project's services and as a vehicle for delivering information. Print-based materials include SEA and project publications and newspapers. Electronic media include slides, audiovisual aids, cassettes, educational television, and films.

Examination of Figure 3.4 shows that, for both 1978 and 1979, the predominant types of media are project and SEA publications. Each of the other media types are used by less than half the projects. Projects, in general, have about four or five of the nine types included in the scale. It appears that after the use of publications, projects will explore various other media forms, and eventually select two or three media types that suit their needs the most, rather than attempt to acquire a wide repertoire of available media linkages.

The secondary comprehensive media linkage scale indicants include closed circuit and commercial television and radio, all of which are generally commercially-produced, rather than SEA-produced, media sources.

Coordinated Linkage

The breakdown of the Coordinated Linkage scale into the various indicant subgroups is shown in Table 3.5. The primary coordinated linkage scale describes the coordination of the linkers utilized by the project with the resource base and with the interorganizational linkages. The ordering of the indicants illustrates a broadening of awareness and usage of linkage networks by linkers and project staff, as well as the development and implementation of linker training. One of the most challenging tasks facing the management of an SCBP project is the integration and coordination of the personal linker agents and resources in order to bring information to the client. The ordering of the indicants in Figure 3.5 shows that this coordination process

TABLE 3.5

COORDINATED LINKAGE ELEMENTS

PRIMARY SCALE

1. The proportion of personal linkage agents who are aware of components of the linkages used by the project i
2. The proportion of personal linkage agents who are aware of linkage services available through the project is
3. The proportion of project staff who are aware of components of the linkage system is
4. The proportion of all personal linkage agents who use the compendium of resources is
5. The usage of the compendium of resources by all personal linkage agents is
6. In responding to client requests, the coupling of personal linkage agents and other resources is
7. The usage of the compendium of linkages by all personal linkage agents is
8. The usage of the compendium of linkages by project staff is
9. Communication flows between linkage agents and the resource base(s) are
10. The usage of training programs for personal linkage agents is
11. The development of training programs for personal linkage agents is

SECONDARY SCALE #1

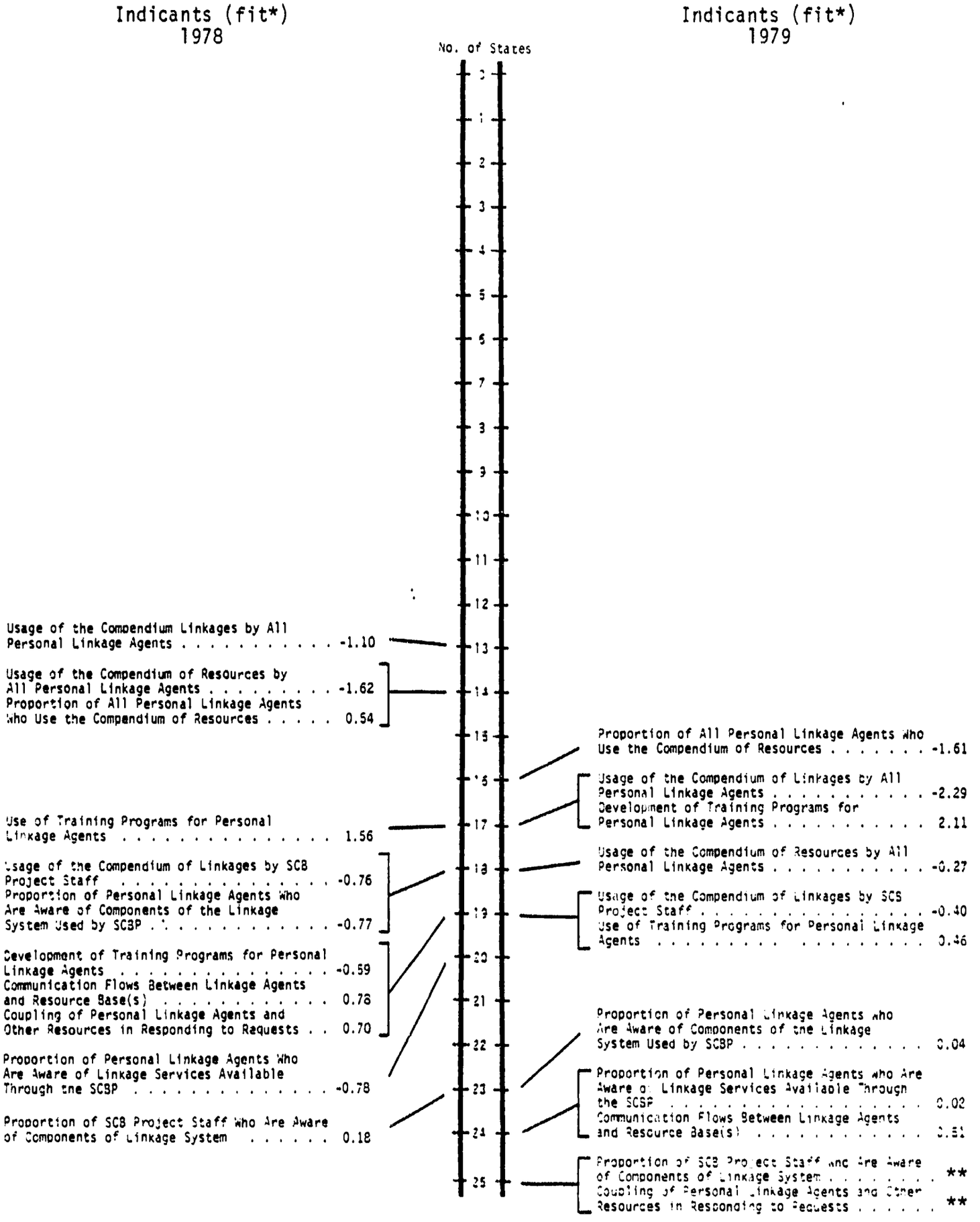
1. The proportion of administrators within the top three levels of the SEA who are aware of the components of the linkage system is
2. The long range planning for coordination of satallite linkages with linkages services of central facilities is
3. The amount that LEA committees determine priorities and information needs of local educators is
4. Recognition of typical problems (e.g., overload, marginality, delivery problems, monitoring problems) is
5. The times a client must contact many sources in order to have his/her information needs met is
6. The amount that duplication of effort is avoided due to a formal referral process is
7. The addressing of typical problems (e.g., overload, marginality, delivery problems, monitoring problems) is

---

Possible responses were: None or Very Low, Low, Moderate, High, Very High

Figure 3.5

COORDINATED LINKAGE ELEMENTS



Scale Reliability = .87
Scale Reliability for 30 Items = .95

Scale Reliability = .78
Scale Reliability for 30 Items = .89

Spearman Rank Correlation = .77

*Total t-test statistic - should be less than 2.0)

requires the following steps: (1) project staff must become aware and knowledgeable of potential resources and linkage components that might be helpful to the project; (2) project staff must then seek out these components and begin to utilize them (i.e., "test them out"); (3) linkers must be trained, with training which ranges from awareness of the available linkages and resources to how to utilize them and how to effect school improvement; and finally, (4) linkers must familiarize themselves with the myriad of resources and linkages, and then proceed to use those that "fit" client needs.

The coordination of linkage activities refers to the extent to which the three major structural components of the system (linker, linkages, resource base) are coordinated in order to provide the most efficient and effective method of allowing the linker to access the system and to bring information to the client. It includes awareness and usage of available linkage networks by project staff and by linkers and the training of linkers, where the training emphasis corresponds to the functions that linkers are expected to fulfill. Table 3.5 also illustrates a broadening of awareness and usage and the increased coordination between the resource base and the linkers. Client requests are responded to through the interaction of resource base personnel and linkers.

While the major coordinated linkage scale includes four stages within the process of coordinating linkers and resources, two aspects that seem to bracket this process are included in the secondary scale. At one end is awareness and planning for linker services; at the other end is the recognition and addressing of typical problems encountered by linkers and duplication of effort.

Institutionalization

The breakdown of the Institutionalization scale into the various indicant subgroups is shown in Table 3.6. The primary institutionalization scale describes the process by which the project's functions and activities are continued after the grant period. The scale includes two major parts. One portion is project-specific, and describes various stages of initial activities in the adoption of the project within the SEA, ranging from awareness, interest, evaluation, trial, and adoption (see Rogers, 1971). The second part describes the gaining of support within and then outside the SEA to insure its permanent status, including agencywide planning for dissemination, on-paper commitments, and increased funding provisions. Examination of the ordering of the indicants in Figure 3.6a and 3.6b suggests that the scale may be viewed as encompassing two separate but related parts. The first part (3.6a), as represented by the elements in the lower range of the scale, is project-specific and parallels Rogers' stages of adoption. This portion of the scale describes initial activities which should lead to institutionalization, such as developing efforts to create awareness and interest in the project and generating a demand for services. The second part (3.6b) of the scale, as represented by the top half of the scale, describes the project's attempts to gain support within and then outside the SEA, namely attempts to insure institutionalization. Probable hallmarks of acceptance of the project include: the gaining of representation in executive team sessions, understanding by personnel both within and outside the SEA of the project's functions, and provisions made for gradual increases of support during the grant period and after the grant period ends.

The second portion of the scale describes the later phases of institutionalization of the dissemination function beyond the identity of the

TABLE 3.6

INSTITUTIONALIZATION

PRIMARY SCALE

1. Planning on an agencywide basis for dissemination is
2. Understanding of a common definition of dissemination on an agencywide basis is
3. The provision for state funding of project activities after the grant period ends is
4. Efforts by the project to stimulate increased demands are
5. Efforts to gain support from clients, potential support groups, and others within the larger organization are
6. The project's "conversational credibility" in the SEA (e.g., project functioning is recognized and interest is shown in it) is
7. The centralization of management of dissemination activities is
8. The provision for gradual increases of state support throughout the project period is
9. Documentation regarding the project (e.g., position papers, role descriptions, standard operation procedures, quality control procedures, rationale) is
10. Budgeting on an agencywide basis for dissemination is
11. State board action on dissemination other than action related to project funding is
12. The consideration of the dissemination function in regular planning activities in the SEA is
13. The mentioning of dissemination in the SEA goals is
14. The contribution of other federal and state funding sources to a coordinated function of dissemination is
15. The amount of state legislation dealing specifically with dissemination is
16. The awareness by those outside the SEA of the functions being performed by the project is
17. Efforts to create awareness among clients, potential support groups, and others within the larger organization are
18. The amount dissemination is mentioned in state board goals is
19. The understanding by those in the SEA of the role behaviors performed by those in the project is
20. Mentioning of the function of dissemination in the state superintendent's annual report is
21. The running of articles related to project activities in agency publications is
22. The extent to which the project was planned by an agencywide group was
23. The participation in executive team sessions, or their equivalent by someone who is closely related to the project and designated as official dissemination representatives is
24. The amount that the project's five year plan is updated as needed and regularly made available is
25. The function of dissemination is located on the organization chart of the SEA
26. There is an agency line item specific to the function of dissemination in the SEA

REMAINING INDICANTS

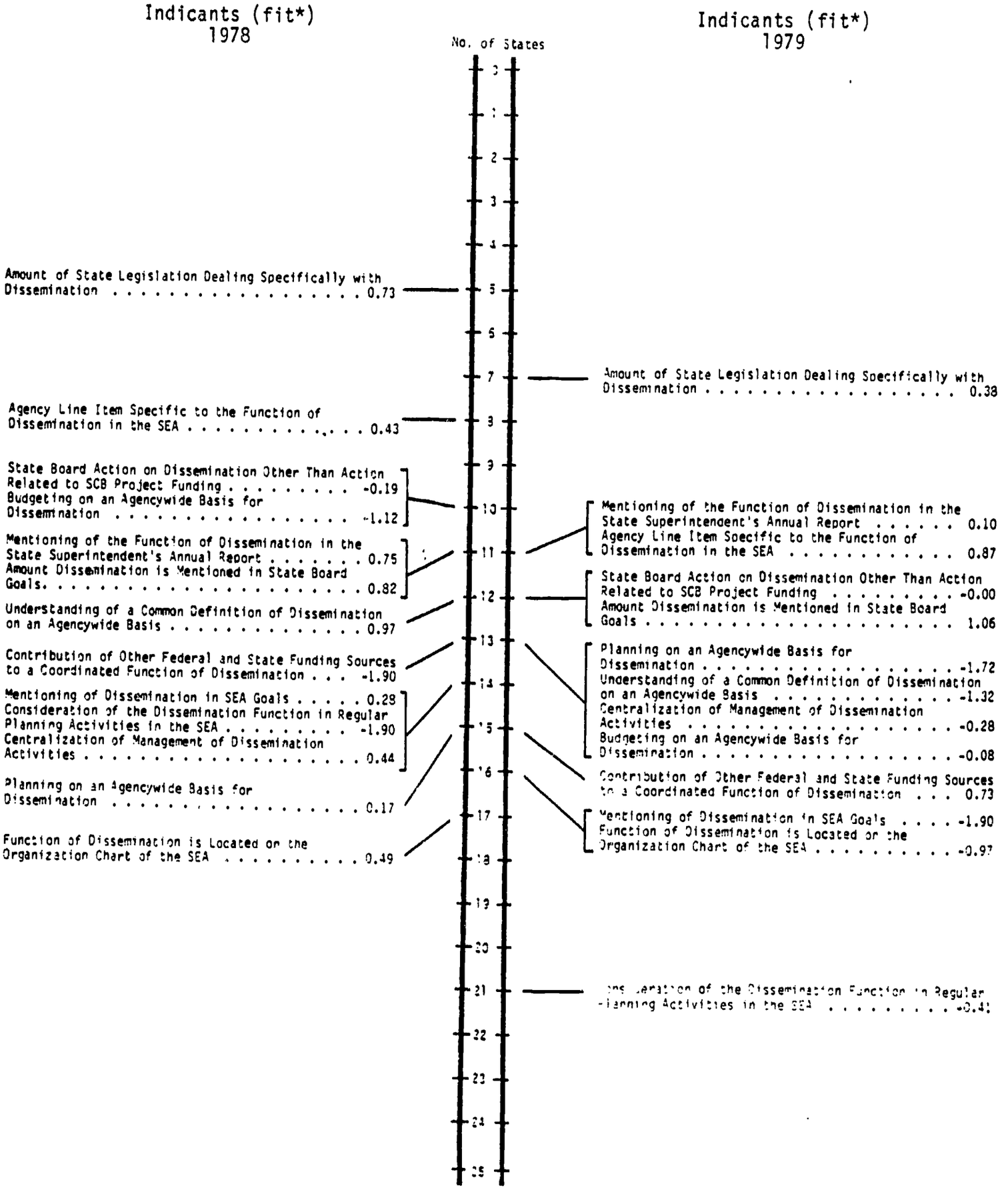
1. The involvement of project staff in preparation of dissemination plans for those federal and state plans requiring dissemination (e.g., 94-142, IV-C, NDN) is
2. The constancy of "titled" roles within the SEA related to dissemination despite personnel changes is
3. The involvement of the project director in responding to the NDN solicitation was
4. The training and recruitment for well defined role positions related to dissemination in the SEA is
5. The acceptance by those outside the SEA of the functions being performed by the project is

---

Possible responses were: Non Existent, Limited, Moderate, Extensive, Very Extensive

Figure 3.6a

KEEPING THE SYSTEM GOING (INSTITUTIONALIZATION, PART 2)



Scale Reliability = .89
Scale Reliability for 30 Items = .90

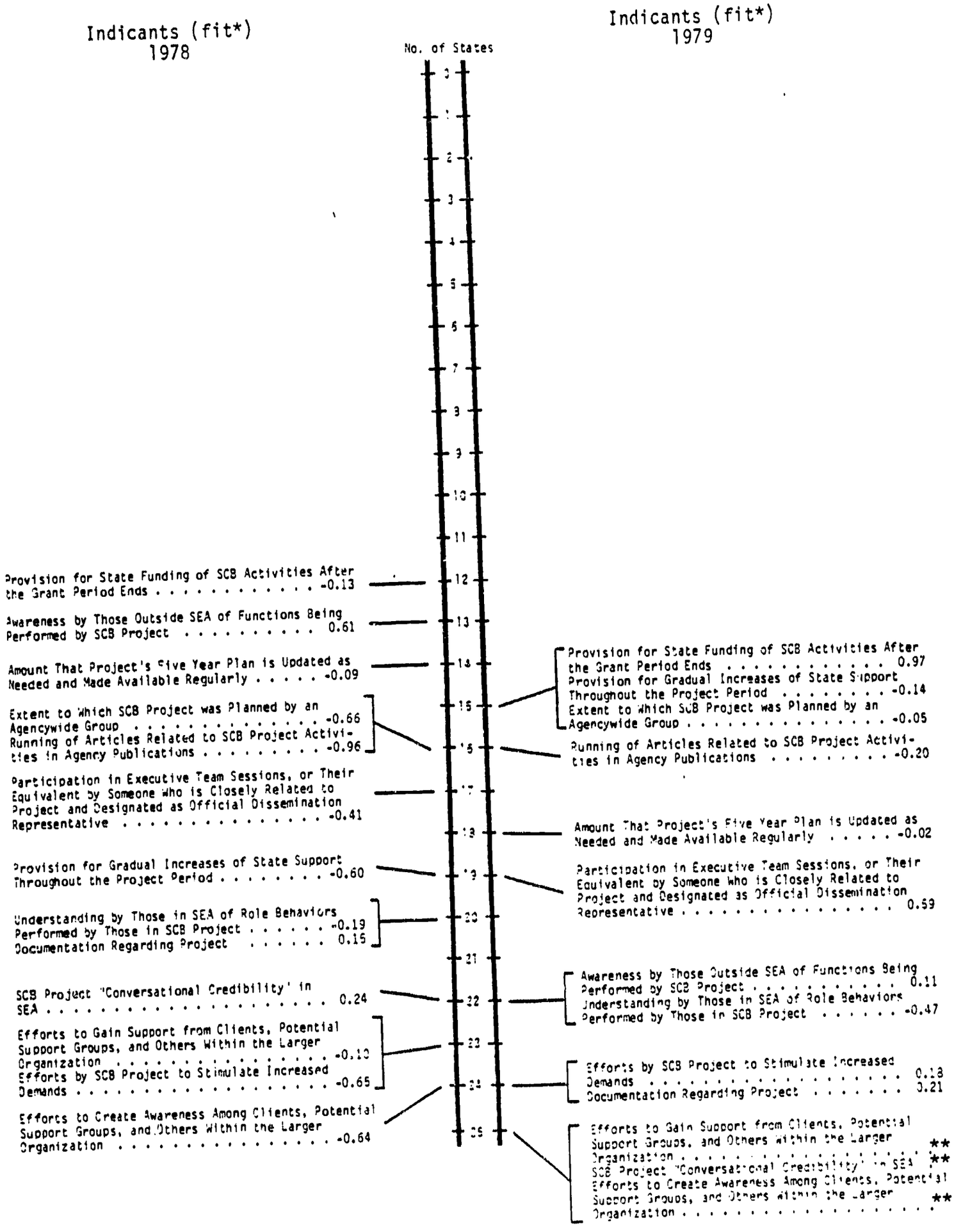
Scale Reliability = .82
Scale Reliability for 30 Items = .86

Spearman Rank Correlation = .85

tal t-test statistic - should be less than 2.0)

Figure 3.6b

GETTING THE SYSTEM GOING (INSTITUTIONALIZATION, PART 1)



Scale Reliability = .89

Scale Reliability = .82

(*Total t-test statistic - should be less than 2.0)

project. Three groups of indicants describe this process: planning for dissemination on an agencywide basis, on-paper commitments to a general dissemination function, and increased funding provisions. The most systematic process involves initial goal statements in the agency, then planning activities that capitalize on project input (in terms of role definitions and experiences), mechanisms for coordinating funding for dissemination, and increasing commitments for future funding. Planning is expanded from planning for the project and its activities to planning for dissemination on an agencywide basis. This may involve a temporary or permanent centralization of management and representation of the project in executive team sessions for planning. There is usually an increase of "on-paper" commitments to a general dissemination function, including goal statements (by the SEA, CSSO, and/or state board), state legislation, location of the dissemination function on the organizational chart, and an agency line item specific to the function of dissemination. Increase in funding occurs as the state specifically budgets for dissemination, utilizes funds from either Federal, state, or other sources, and makes provisions for support of project activities both during and after the grant period.

Thus, the institutionalization scale describes two phases of the institutionalization process: adoption of the project by the SEA, and institutionalization of the dissemination function by the agency. Implementation, that is, how the project builds dissemination capacity and what that capacity looks like are captured by the previously described scales.

The remaining indicants include two items that assessed the involvement of project staff in other dissemination efforts and two that involved the extent to which dissemination roles are well-defined. One possibility for their exclusion from the scale is that they are not necessary activities leading towards institutionalization.

CAPACITY BUILDING AS A DEVELOPMENTAL PROCESS

In Section 3, we discussed the content of the scales in some detail. In those discussions we noted that there appeared to be a pattern to the activities expressed in the scales. These patterns, such as the early acquisition of materials most easily obtainable to finally obtaining materials that are more difficult to develop, seemed to indicate a sequence of behaviors to which we suggested a developmental paradigm. However, our interpretation of the content of the scales was not sufficient for verifying a developmental sequence. Therefore one of the purposes of this section is to examine this developmental assumption.

In the Interim Report, we presented scales based upon data collected at one point in time (i.e., data were from a cross-sectional field collection effort). This enabled us to describe projects and dissemination systems, but not how they grow or develop over time. However, because data were collected from projects and SEAs that had been in the State Capacity Building Program for different numbers of years, we considered the possibility that these projects were in different phases in developing their dissemination capacity. In fact, moderate correlations had been found between the number of years in the Program and most of the scale scores. Therefore, we considered the possibility that the scales could be used to describe a developmental sequence that states go through in the course of building capacity. One of our major objectives is, of course, to trace the development of systems over time. If we can verify that the scales reflect the developmental patterns or stages within various aspects of a system, then we have a general developmental pattern and a means for assessing individual or group differences. We can do

this by comparing scale scores for states and groups of states on the basis of cohort, time of measurement, and years in the program, and on contextual factors including initial dissemination capacity, size, and so forth.

While a developmental interpretation is relatively easy to apply to the content of the scales, such an interpretation is also fraught with serious problems. Consider, for example, the Comprehensive Resource Base Scale. It was intuitively appealing to consider the development of the resource base as proceeding through the acquisition of nationally-available data bases (e.g., ERIC), moving through the stages of collecting materials which are more difficult to track down and obtain (e.g., local exemplary program files) to the development of materials which the project may have to develop on its own (e.g., a human resources file). However, there are real problems in making such conclusions on the basis of these cross-sectional data, since development might not follow that pattern over time. In order to more accurately assess the actual pattern of development, projects needed to be examined across time periods. We needed to verify whether the ordering of the indicants represents a developmental progression in the building of capacity within particular facets, or whether the scales are a cumulative frequency that measures the quantity and rarity of each indicant. For example, the higher up an indicant is, the rarer it might be, rather than representing a characteristic that is added to a system at a later point in a project's life. Interestingly, while the scales measuring the comprehensiveness of resources and of program linkages are more demonstrably valid, they are also more likely to be interpreted as measures of rarity. The scales measuring coordination and institutionalization seem more likely to show progression in a developmental sequence.

In order for the scales to be measuring a developmental process, a state's position on the scale at one point in time should be predictable from a knowledge of that state's position on the scale at a previous point in time. A necessary condition for a scale to be developmental is that the scale be hierarchical. A scale is hierarchical if a positive score on a particular indicant is accompanied by a high probability of positive scores on all of the other indicants below it on the scale. Conversely, a negative score on a particular indicant is accompanied by a high probability of negative scores on all of the other indicants above it on the scale. As a consequence, the indicants on a hierarchical scale are ordered by their degree of difficulty.[4.1] In order to utilize the scales to measure growth, it is necessary to verify whether the scales are just hierarchical and measure the quantity and rarity of each indicant's occurrence, or whether the ordering of the indicants is not only a hierarchical but a developmental progression in the building of capacity within a particular facet. If the former case is true, then this means that the higher an indicant is on the scale, the rarer it is, as opposed to the indicant representing a characteristic that is added to a facet at a later time in the building of capacity.

As a further clarification of the distinction between hierarchical and developmental scales, consider two measurements that we frequently make of

---

4.1 As introduced in Section 3, the difficulty of an indicant is directly related to how many states had negative responses with respect to that indicant. The more states that respond negatively to an indicant, the more difficult it presumably is for any given state to respond positively to an indicant. What is meant by a positive response will of course vary from scale to scale. For the Comprehensive Resource Base Scale, a positive response indicates that a particular resource is available through the dissemination system. For Institutionalization, a positive response signifies that an activity associated with institutionalization occurs frequently or extensively in the state. In summary, an indicant with a higher degree of difficulty is associated with fewer positive responses; an indicant with a lesser degree of difficulty is associated with more positive responses.

ourselves: body temperature and height. The fever thermometer scale is hierarchical; when the mercury is at 98.6, it is also at every degree marking below 98.6; by the same token, if the mercury is not at 98.7, then it is not at any degree marking above 98.7. The fever thermometer scale, however, is certainly not developmental, as an individual's temperature is a function not of the individual's age but rather of various life events, such as being exposed to an influenza virus. Our measuring tape scale, on the other hand, is not only hierarchical but is also developmental. The height scale is hierarchical because if one is taller than 6 feet, one is also taller than 5 feet 11 inches, 5 feet 10 inches, etc. However, it also measures a developmental sequence because from the time we are born at least generally up through our teenage years, the process of maturing includes growing taller.

In Section 3, we established that the indicants within a scale form an order which is consistent not only across states, but also across two points in time. Since this order was consistent, we can therefore say that the indicants for each scale form a hierarchy, and that hierarchy is invariant across time. This means that the scales are appropriate not only as devices with which the states can be measured and ordered with respect to how much of each dissemination system facet they possess, but also with which the growth of the states can be measured with respect to each of the facets. In the remainder of this section, we consider whether the indicants for each scale constitute a developmental sequence.

In order to assess whether the scales reflect a developmental paradigm, the scale scores should show a consistent relationship with age. "Age" in this context refers to the number of years that a state has been in the capacity building program. Consequently, it is necessary to have measures of

the states at at least two points in time in order to perform this assessment. When measurement is available from only one point in time, we can only perform a cross-sectional analysis. Such an analysis allows us to test whether there are differences in scale scores for states from different cohorts (i.e., of different project ages) measured at the same time, but such differences will be completely confounded with any cohort differences that may exist. Such cohort differences will be systematic in nature (e.g., the states in one cohort may have had substantially greater initial dissemination capacity than the states in another cohort), but will be completely undetectable in a cross-sectional design. A second factor that will also be completely confounded with growth will be the point in time at which the development of the system is measured. For example, if our measures were taken in a year following a period in which the dissemination community became aware of the importance of building a comprehensive resource base, then all of the states might have immediately acquired as many resources as possible. The sudden emphasis on acquiring resources might have completely altered any potential differences in number of resources between the states as a factor of the number of years that each state had been in the capacity building program.

In order to investigate the extent to which the scales reflect a developmental paradigm (i.e., a relationship between program effects and number of years in the program) as opposed to the possible confounding of effects of cohort and/or time of measurement, a paradigm suggested by Schaie (1965, 1973) is particularly useful. Schaie's paradigm permits the investigation of the development of systems by considering the degree to which that development is affected by three components: years in the program, the point at

which capacity is measured, and the cohort to which a state belongs. Table 4.1 graphically depicts the paradigm, including cohorts, project ages, and measurement points, and will serve as a heuristic device in understanding the discussion which follows.

Table 4.1
Analytic Paradigm

Cohort	Ages				
I	0	1	2	3	4
II	-	0	1	2	3
III	-	-	0	1	2
IV	-	-	-	0	1
V	-	-	-	-	0
Non-SCBP	-	-	-	-	0
Time of Measurement	1975	1976	1977	1978	1979

---Cohort-Sequential Method Data
---Time Sequential Method Data

The first of these procedures which Schaie refers to as the cohort-sequential method involves examining the data enclosed by the dotted line in Table 4.1. This data configuration requires information on projects drawn from two cohorts that have been in the program for either two or three years. Using this data, we can assess the effect of years in the program, and see if this effect remains constant across two different cohorts. In order to assess the average age change over one year starting at age two, as sampled in cohorts I and II, we use the formula:

$$\text{AVERAGE AGE DIFFERENCE} = \frac{[(\text{Age 3 Coh I}) - (\text{Age 2 Coh II})] + [(\text{Age 3 Coh II}) - (\text{Age 2 Coh III})]}{2}$$

To assess whether this effect remains constant over two different cohorts, we use the formula.

$$\text{COHORT DIFFERENCE} = [(\text{Age 3 Coh I}) - (\text{Age 2 Coh I})] - [(\text{Age 3 Coh II}) - (\text{Age 2 Coh II})]$$

If we should find that the effect of age is different for Cohort I than it is for Cohort II (i.e., the cohort difference is not equal to zero), then the usefulness of the results will be questionable. This approach, however, requires us to assume that the time of measurement has no effect. If this assumption is not met, then the differences that we might ascribe to years in the program could instead be attributed to time of measurement effects. If, for example, a breakthrough occurred between 1977 and 1978 and another breakthrough occurred between 1978 and 1979, then the growth that we might see would be attributable to exposure to these breakthroughs, and not to the effect of the program.

The second of these procedures which Schaie refers to as the time-sequential method involves examining the data enclosed by the solid line in Table 4.1. In this data configuration, we have information on projects that have been in the program for two years or three years drawn at two points in time. Consequently, we can assess the effect of years in the program and see if this effect remains constant across two different points in time. In order to assess the net age change over one year starting at age 2, as sampled at Measurement Times 1978 and 1979, we use the formula:

$$\text{AVERAGE AGE DIFFERENCE} = \frac{[(\text{Age 3 Coh I}) - (\text{Age 2 Coh II})] + [(\text{Age 3 Coh II}) - (\text{Age 2 Coh III})]}{2}$$

To assess whether this effect remains constant over two different points in time, we use the formula:

$$\text{TIME DIFFERENCE} = [(\text{Age 3 Coh I}) - (\text{Age 2 Coh II})] - [(\text{Age 3 Coh II}) - (\text{Age 2 Coh III})]$$

If we should find that the effect of age is different in 1978 from what it is in 1979 (i.e., the time difference is not equal to zero), then the usefulness of the results will be questionable. This approach, however, requires us to

assume that the cohort membership has no effect. If this assumption is not met, then the differences that we might ascribe to years in the program could instead be attributed to cohort effects.

Each method, if used alone, confounds growth with either cohort membership or time of measurement. Utilizing both methods to assess the effects of growth allows us to "zero in" on the effect of years in the program. Specifically, if the differences that we observe in both methods are in fact due to increasing years in the program, and not to cohort membership or time of measurement, then we should observe exactly the same relationship between the scale scores and number of years in the program regardless of which method is used. To the extent that confounding does exist, either due to cohort membership or time of measurement, then combining these two methods should allow us to assess the strength of this confounding and to temper our conclusions about the effects of years in the program accordingly.

Unfortunately, we do not have the complete information needed to implement the cohort-sequential method. Information was collected only in 1978 and 1979, so the information needed for the cohort-sequential method from 1977 is unavailable to us. Consequently, our estimate of average age difference has to be based only on one pair of scores rather than two. For example, the effect of age change over one year starting at age two can be estimated only by [(Age 3 Coh II) - (Age 2 Coh II)]. Furthermore, we are unable to directly compare the effects of age across two different cohorts to insure that the effects of age are not confounded with cohort membership. To the extent, however, that our estimates of the age effect tend to converge, then TIME DIFFERENCE, from the time sequential method, will be zero, and we can assume that the COHORT DIFFERENCE, from the cohort sequential model, would also probably have been zero.

From our data, we can estimate the effects of net age change over one year for projects starting either at age one or at age two. The information that we will use is described in Tables 4.2a through 4.2f.

Table 4.2a
Comprehensive Resource Base

Cohort	Age	Mean	SD	n	Age	Mean	SD	n
I	3	12.8	.83	7	4	11.28	1.38	7
II	2	11.7	2.55	13	3	11.8	1.70	13
III	1	11.0	2.76	5	2	11.0	2.83	5
IV	0	--	--	-	1	8.3	2.08	3
V	0	--	--	-	0	6.8	2.28	5
Non-SCBP	0	--	--	-	0	9.0	2.0	2
Time of Measurement		1978				1979		

Table 4.2b
Coordinated Resource Base

Cohort	Age	Mean	SD	n	Age	Mean	SD	n
I	3	4.86	1.73	7	4	5.0	1.51	7
II	2	5.61	1.55	13	3	5.61	1.59	13
III	1	4.6	1.85	5	2	5.6	1.50	5
IV	0	--	--	-	1	5.0	1.0	3
V	0	--	--	-	0	2.8	1.30	5
Non-SCBP	0	--	--	-	0	2.5	1.5	2
Time of Measurement		1978				1979		

Table 4.2c
Comprehensive Program Linkage

Cohort	Age	Mean	SD	n	Age	Mean	SD	n
I	3	7.71	1.48	7	4	7.28	1.66	7
II	2	5.31	2.94	13	3	6.69	2.27	13
III	1	6.4	2.58	5	2	7.2	2.64	5
IV	0	--	--	-	1	3.0	3.61	3
V	0	--	--	-	0	3.0	2.82	5
Non-SCBP	0	--	--	-	0	8.50	.5	2
Time of Measurement		1978				1979		

Table 4.2d
Comprehensive Media Linkage

Cohort	Age	Mean	SD	n	Age	Mean	SD	n
I	3	4.0	2.0	7	4	4.71	2.18	7
II	2	3.38	2.27	13	3	3.54	1.86	13
III	1	3.0	3.52	5	2	4.0	1.78	5
IV	0	--	--	-	1	2.0	2.0	3
V	0	--	--	-	0	3.0	3.32	5
Non-SCBP	0	--	--	-	0	--	--	2
Time of Measurement	1978				1979			

Table 4.2e
Coordinated Linkage

Cohort	Age	Mean	SD	n	Age	Mean	SD	n
I	3	7.86	3.83	7	4	9.42	1.18	7
II	2	7.77	2.75	13	3	8.46	2.50	13
III	1	7.6	3.38	5	2	10.2	1.6	5
IV	0	--	--	-	1	5.0	5.0	3
V	0	--	--	-	0	3.4	3.51	5
Non-SCBP	0	--	--	-	0	--	--	2
Time of Measurement	1978				1979			

Table 4.2f
Institutionalization

Cohort	Age	Mean	SD	n	Age	Mean	SD	n
I	3	17.7	8.68	7	4	18.43	5.92	7
II	2	15.5	4.46	13	3	16.07	3.85	13
III	1	13.6	2.92	5	2	20.0	2.83	5
IV	0	--	--	-	1	13.3	9.5	3
V	0	--	--	-	0	6.2	3.70	5
Non-SCBP	0	--	--	-	0	3.5	2.5	2
Time of Measurement	1978				1979			

It might be noted that if information had been collected from the Cohort IV states in 1978, then it would have also been possible to estimate the effect of one year's growth for cohorts beginning at age zero. Without the 1978 Cohort IV information, the possible effects of growth between ages zero and one are completely confounded with cohort effects and time of measurement effects.

Also included in Tables 4.2a through 4.2f are the number of projects in each cell. Of the 29 projects in Cohorts I through III, 25 supplied information on the capacity building indicants questionnaire for both 1978 and 1979. Three of the four cohort IV states and five of the ten Cohort V states supplied information in 1979. Finally, we have information from two states that are not currently in the SCBP program.

The impact of growth on scale scores is presented graphically and includes all six groups (Cohorts I through V and Non-SCBP) for 1978 and 1979. Figures 4.1 through 4.6 each present the relationship between scale scores and number of years in the program. The mean scale scores for each cohort and time of measurement cell are presented as a function of years in the program. The mean scale scores from 1978 are represented by squares and joined together with line segments. The 1979 mean scale scores are represented by circles, and are also joined together by line segments for Cohorts I through V, while the non-SCBP average score has been left as a distinct point.

Given the small sample size and the resultant loss of statistical power, we decided that it would be more appropriate to examine the graphs in Figures 4.1 through 4.6 in order to assess the developmental paradigm than to perform a series of statistical tests. In examining Figures 4.1 through 4.6, several considerations should be kept in mind. First, if the position on the scale

is solely a function of age (i.e., not a function of cohort or time of measurement differences), then each pair of points for years one, two, and three should lie extremely close together (e.g., the 1978 mean scores for cohort II should be close to the 1979 mean score for cohort III). To the extent that the line segments are superimposed on one another, the developmental hypothesis is supported and the cohort-sequential method and time-sequential method have converged. If the lines are not superimposed, then there are several other patterns that we should look for in an effort to understand the factors that might influence position on the scale.

If the lines are more or less parallel to one another, then there are no time difference effects; i.e., whether the measurement was made in 1978 or 1979 has no impact on position on the scale. Position on the scale, therefore, is a function of some combination of age and cohort. Another pattern to look for when the lines are not superimposed is whether shifting the 1978 points one year to the right will cause any of the points to be superimposed. Should a point be superimposed, the positioning of the projects on the scale is almost completely a function of cohort differences. This is because the two overlapping points will represent the same group of projects with one point being the 1978 average score and the other point being the 1979 average score. If no growth occurred between 1978 and 1979 for the same group of projects, then it seems likely that differences between the project averages in 1978 or in 1979 were attributable to differences between cohorts.

The average scores for the various groups on the comprehensive resource base scale are presented in Figure 4.1. The first impression that this figure conveys is that comprehensiveness of the resource base is a constantly increasing function of age, at least up through three years. Whether the

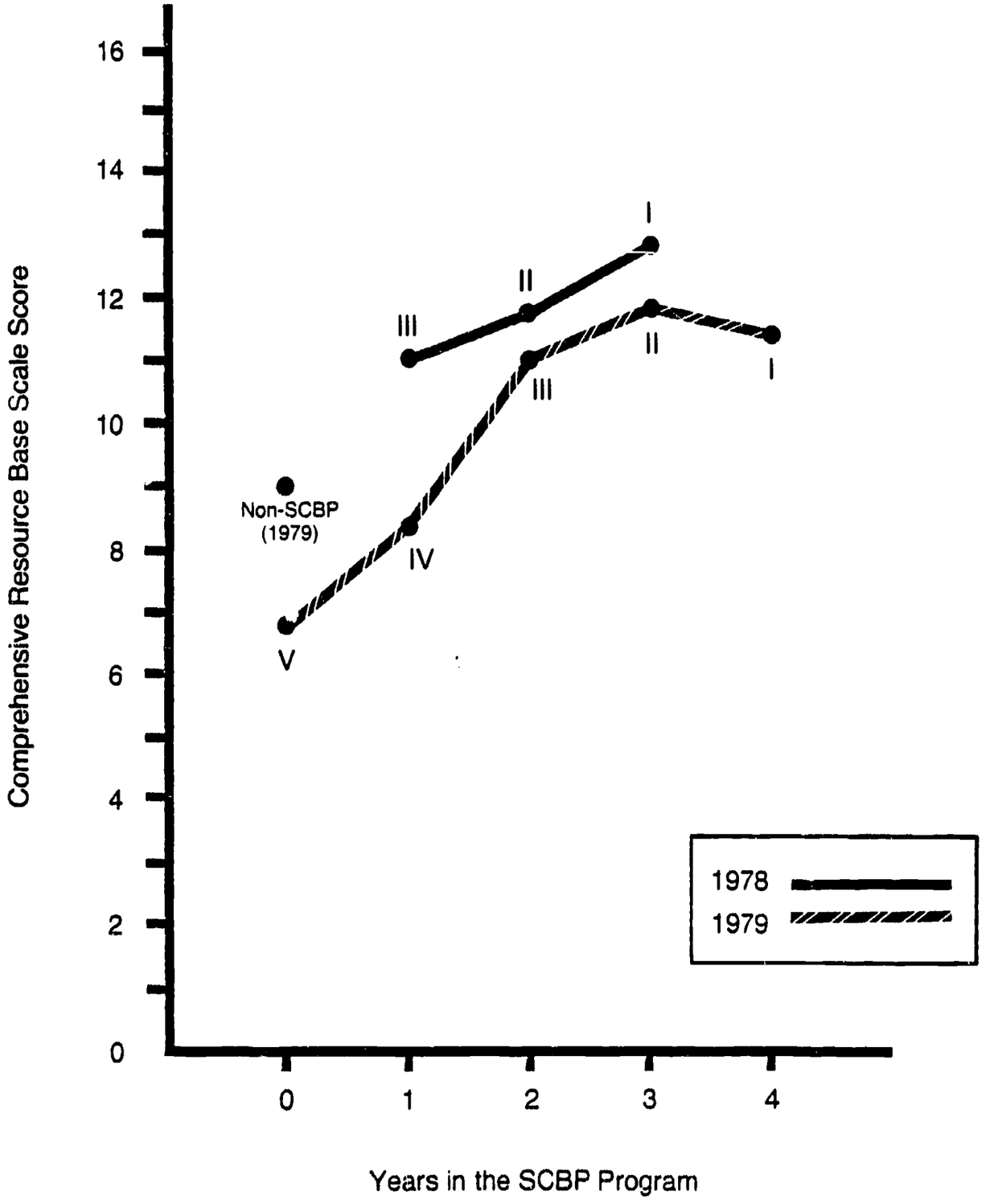


Figure 4.1 Comprehensive Resource Base

dropoff from three years to four years is truly a function of age or is unique to Cohort I projects cannot be tested from this information. Since the 1978 and 1979 lines are not superimposed, it would appear that age cannot be considered the sole determinant of scale position. Upon further examination, in fact, it may well be the case that age is of little importance at least in the period of one to three years. Moving the 1978 line one year to the right superimposes not only the two Cohort II points, but also the two Cohort III points. In other words, from 1978 to 1979, neither Cohort II nor Cohort I changed on the comprehensive resource base scale. Thus, based on this information, the most likely explanation is that the differences between the project groups are attributable not to age differences but primarily to cohort differences.

The average scores for the various groups with respect to the coordinated resource base scale are presented in Figure 4.2. As with the comprehensive resource base scale, the first impression that this figure provides is that coordination is also a monotonically increasing function of age up through three years, with a dropoff from the third to the fourth year. Further examination indicates, however, that two different processes seem to be at work. The position of projects at one and two years of age on this scale would seem to be entirely a function of age. All three estimates for the effect of growing from one year to two years seem to have converged on approximately one unit (approximately 4.8 to 5.6). This is clear support for the developmental metaphor. On the other hand, for both Cohorts I and II, there was virtually no growth from 1978 to 1979. It would appear that beyond two years of age, projects remain at the same position on the scale. The logical inconsistency in this argument, though, is that both of the Cohort I

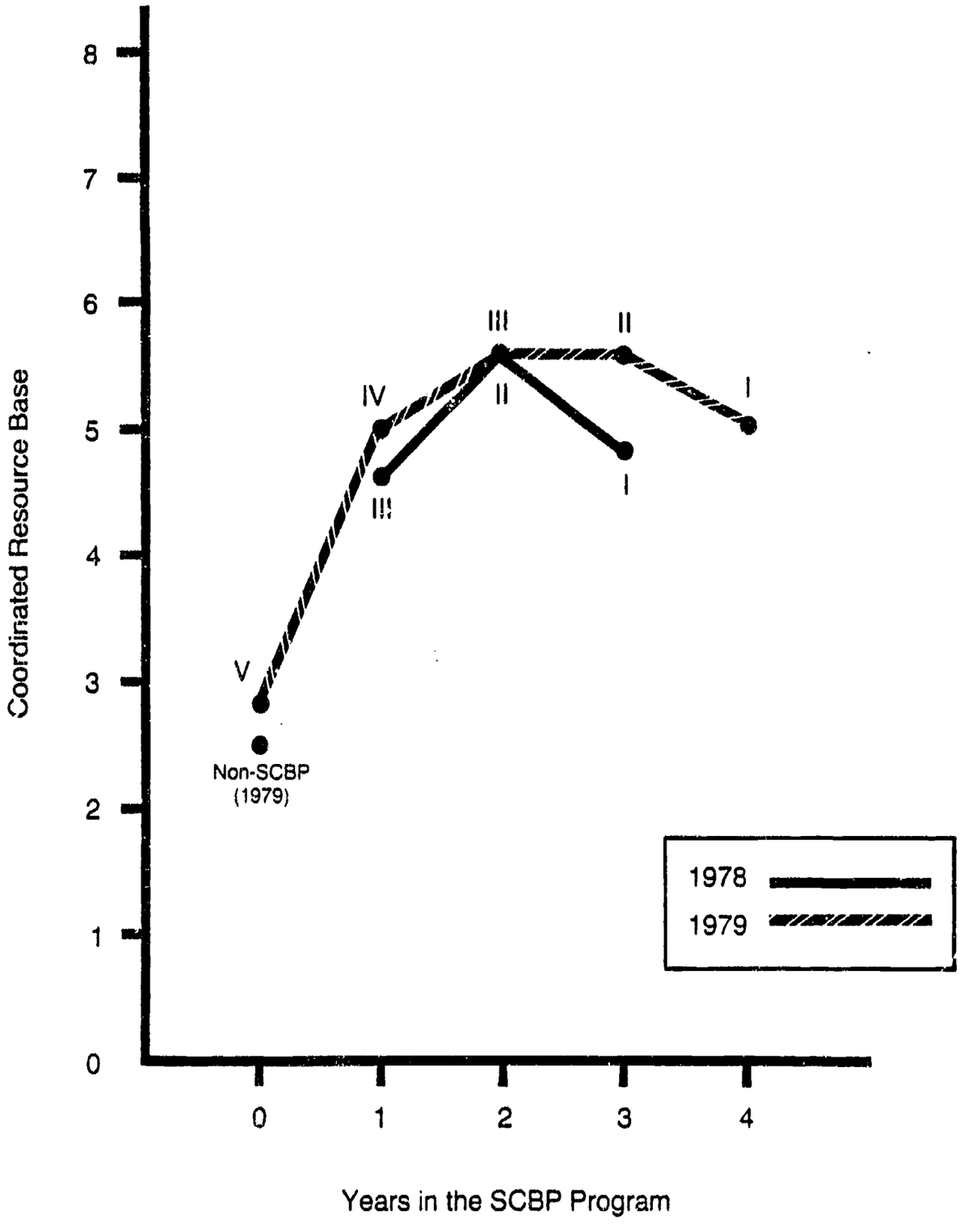


Figure 4.2 Coordinated Resource Base

points should therefore be at approximately 5.6, since that was the point that both Cohorts II and III reached at two years of age. More information than is currently available, especially the position of the Cohort I projects on this scale prior to 1978, is needed to resolve this inconsistency.

The group average scores for the comprehensive program linkage elements scale are presented in Figure 4.3. The validity of the developmental metaphor for this scale immediately appears to be challenged by the configuration points. For both 1978 and 1979, Cohort I projects have the highest average scale score, followed by Cohort III projects followed by Cohort II projects. If we move the 1978 line one year to the right, the Cohort I points are extremely close, while the 1978 Cohorts II and III points both lie approximately the same distance below the corresponding 1979 points. This would seem to suggest that there is a mixture of cohort and age effects operating here. Clearly, the Cohort I states did not show any growth. For the Cohort II and III states, the fact that the Cohort II states scored lower than the Cohort III states also suggests the presence of a cohort effect; however, both the Cohort II and Cohort III states each showed roughly the same amount of growth from 1978 and 1979 (approximately .5), thus suggesting that while cohort differences may have had an effect with respect to initial position, an additional year in the program did have an impact on increased scale scores, at least for year 1 to year 2 and for year 2 to year 3. This is only a tentative hypothesis, however, and would require at least another year's worth of information for its validity to be supported.

The group average scores for the comprehensive media linkage scale are displayed in Figure 4.4. With the exception of the 1979 Cohort II average score, the impression that this figure presents is of a constantly increasing

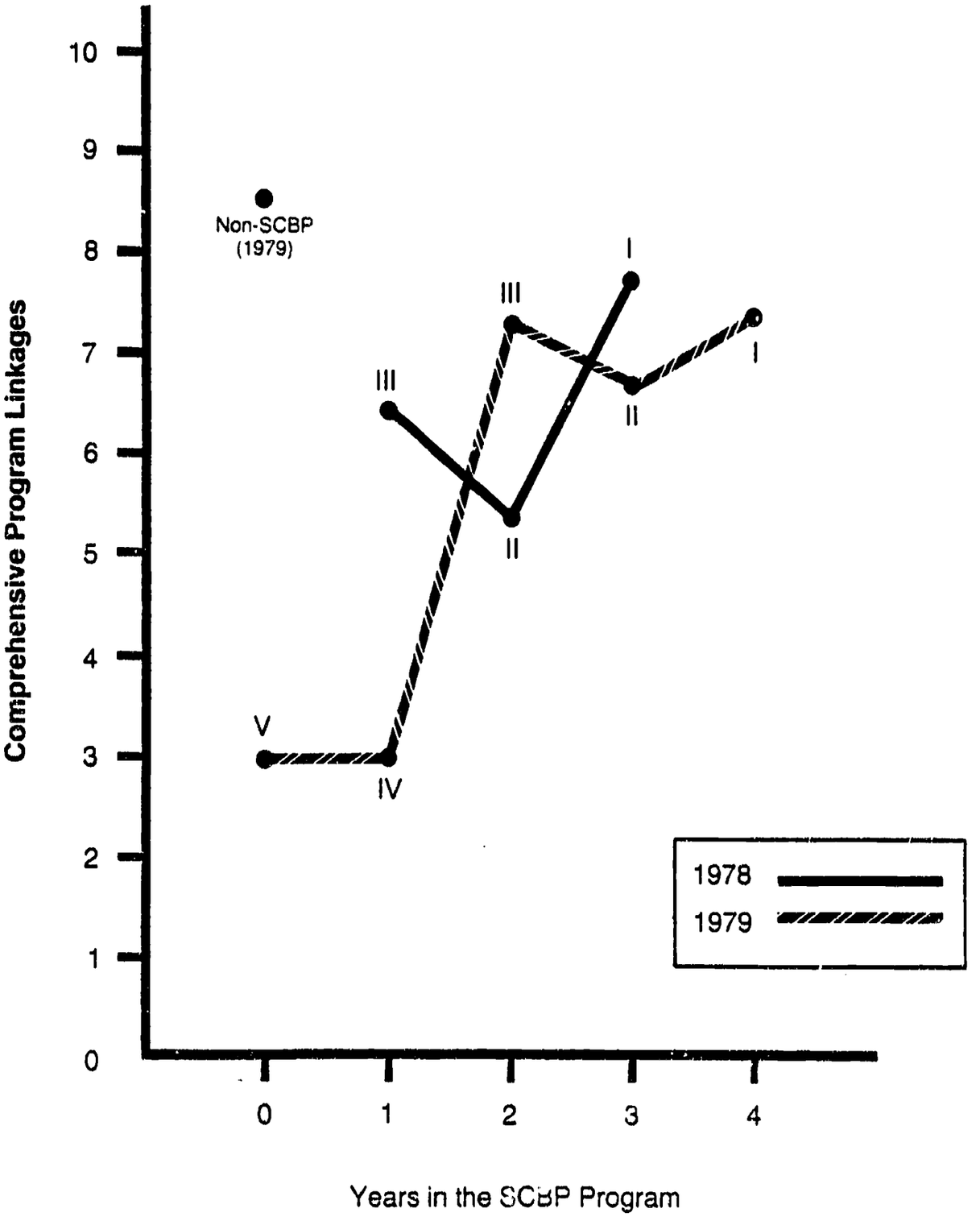


Figure 4.3 Comprehensive Program Linkage

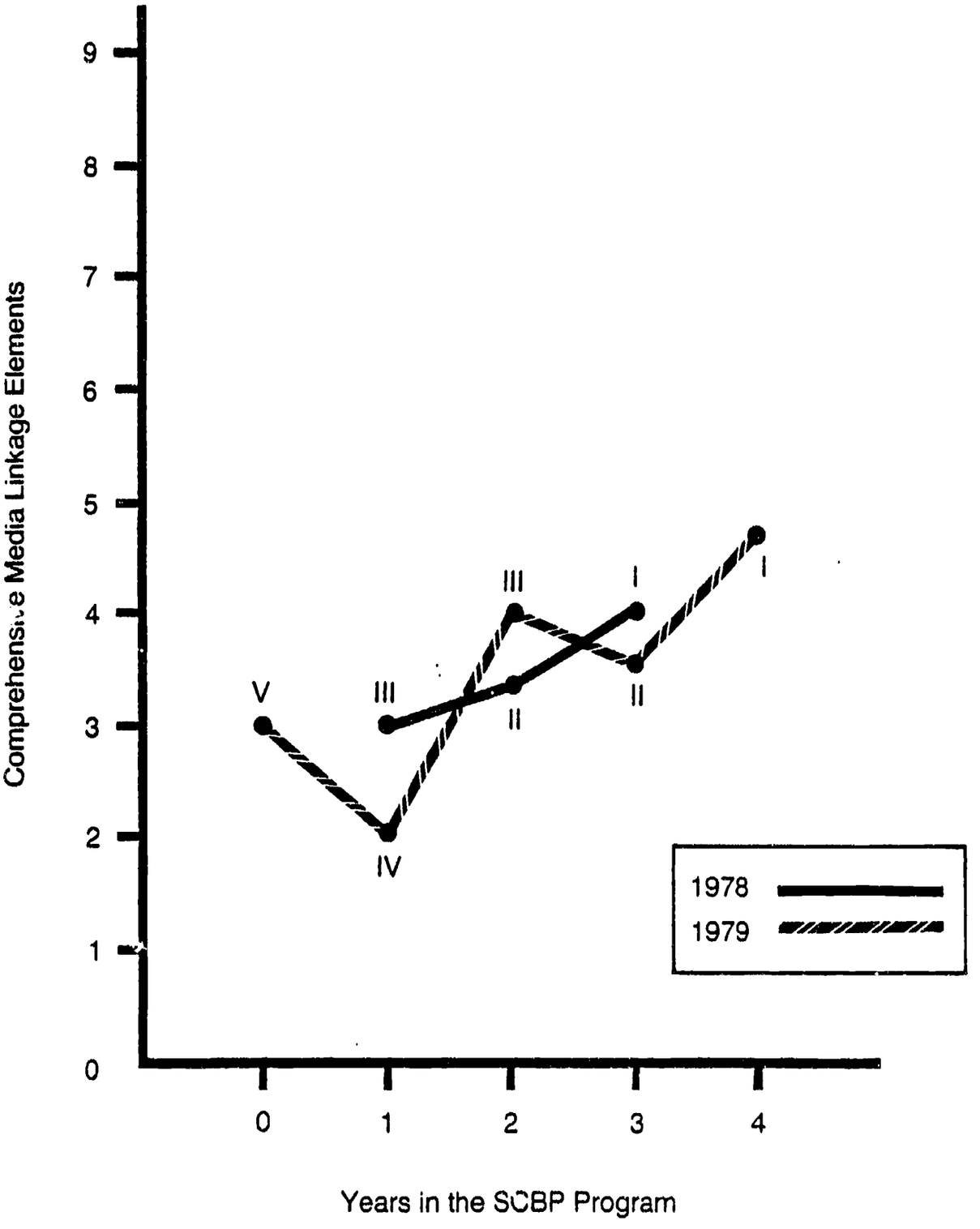


Figure 4.4 Comprehensive Media Linkage

relationship between average scale score and number of years in the program. It should be noted in Table 4.2d that each of the group average scores has a substantial standard deviation (close to or in excess of 2) associated with it. Given that the difference between the pairs of points at years 2 and 3 is less than .5, it is difficult to determine whether the lack of overlap between the 1978 and 1979 lines is due to the confounding effects of cohort and/or time of measurement, or is simply a function of measurement error. Regardless, the fact that the standard deviations are so large relative to the range of the scale renders any interpretation of the comprehensive media linkage elements data virtually meaningless.

The average group scores for the coordinated linkage elements scale are presented in Figure 4.5. The 1978 line shows practically no differentiation between cohorts, while the 1979 line shows substantial but erratic growth. Given such a pattern, it is impossible to even guess at the relative contributions of age, cohort, and time of measurement without additional information.

The group average scores for the institutionalization scale are presented in Figure 4.6. In general, there would appear to be substantial cohort effects, although there are some indications of an age effect in the growth from one to two years. If the 1978 line is moved over one year to the right, the Cohort I points are extremely close given the size of the standard deviations in Table 4.2f, as are the Cohort II points. In other words, there was little, if any, growth with respect to institutionalization for Cohorts I and II from 1978 to 1979. The substantial differences between the 1978 and 1979 lines with respect to growth between one and two years further suggests the presence of a cohort effect.

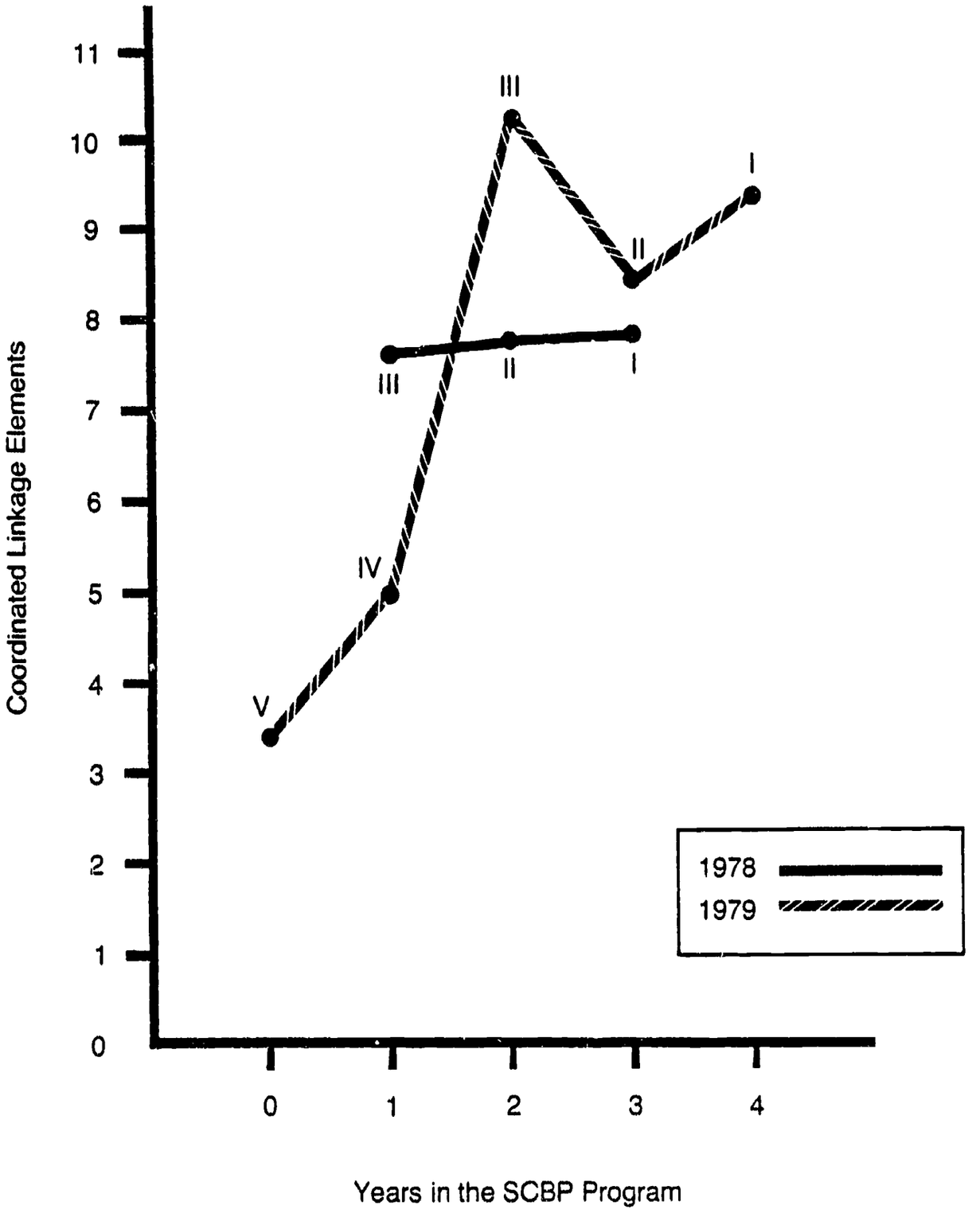


Figure 4.5 Coordinated Linkage

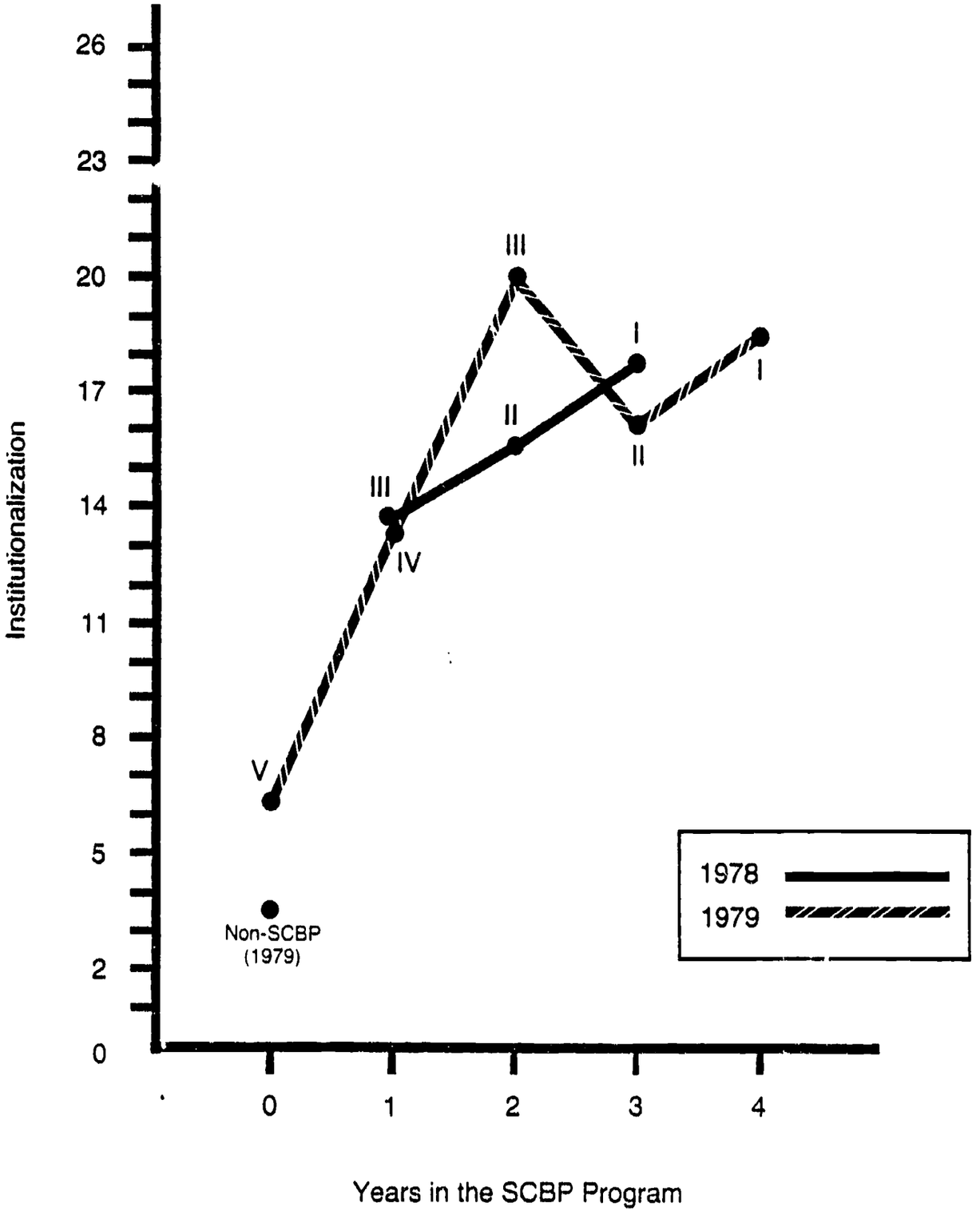


Figure 4.6 Institutionalization

It might be noted that the Cohort V group averages are consistently below the averages of all of the other groups. Whether this represents the distinction between being in the project for zero years as opposed to one year or more cannot be separated from the possibility that the Cohort V states, being the last to enter the program, were states which had not engaged in many dissemination efforts and therefore had a lower initial status than all of the other states. With the exception of the comprehensive program linkage elements scale, the Cohort V and Non-SCBP averages tended to be close together, indicating that the differences between states that have just entered the program and unfunded states are minimal.

5

CONCLUSIONS

While the previous chapters of this substudy have focused on such details as how the facets were conceptualized, how the facets were turned into scales, and how the developmental nature of the facets was assessed, the purpose of this chapter is to "step back" from such detail and assess the overall utility of these scales. Such an assessment of utility must necessarily be based on the extent to which information regarding the scale scores of the dissemination system can be used for purposes of program planning and improvement.

To begin with, we believe that the development of the scales represents a significant methodological advance in understanding and conceptualizing the organizational attributes of a dissemination system. Typically, the organization of a dissemination system has been described in qualitative rather than quantitative terms. A case study approach has frequently been used to generate a great deal of descriptive information regarding the system, which, when combined with information from other systems, could be combed for points of similarity and dissimilarity. Such an approach, while rich in detail, nonetheless makes it difficult for the individual to develop some sort of unified framework into which all of the systems can be placed. Such a framework permits a much clearer perception of system similarities and differences, and provide a much more concise description of each system. With such a description in hand, one can then begin to assess the impact of the environment and the implementation of the SCBP program on the dissemination system's characteristics, as well as to assess the degree to which the

dissemination system characteristics result in effective utilization of the system, which in turn leads to better and better instruction.

The legitimacy of the scales as instruments through which dissemination system attributes can be measured was established in Chapter 3. This legitimacy is based upon the observation that the ordering of the indicants within each scale was consistent across the states, thus demonstrating that all of the indicants within a scale seem to be measuring the same thing. Furthermore, the ordering of the indicants remained generally invariant across two administrations (fall of 1978 and fall of 1979) of the data collection instrument. The indicants within each scale, formed a hierarchy, where a positive response with respect to any indicant was accompanied by a high probability that all of the other indicants lying lower on the scale would also be associated with positive responses. By the same token, a negative response with respect to an indicant was accompanied by a high probability that all of the indicants lying higher on the scale would be associated with negative responses.

The descriptive power of such a set of scales is enormous. Rather than having to characterize a dissemination system through the atomistic approach of listing a number of details regarding that system, it is possible instead to adopt a more holistic approach, synthesizing all of the detailed information into a few scale scores. Knowledge of where a dissemination system lies on a scale automatically carries with it knowledge regarding the dissemination system with respect to a number of details. If we regard the indicants as details, then the scale score infers which attributes the system is likely to possess, as well as which attributes the system is not likely to possess. In short, this substudy has revealed the presence of order among what might otherwise be regarded as a descriptive cacaphony of dissemination system attributes.

In addition to the succinct synthesis of a great deal of detailed information, the fact that the structure of the scales remains invariant both across states and across time has two important implications for how the scales can be utilized. First, states can all be ordered with respect to their score on each scale; in other words, the scores that two states receive on a scale can legitimately be used for purposes of comparison, and conclusions can be drawn about one state being higher or lower or about the same as another state. As will be discussed shortly, some care needs to be exercised in drawing such conclusions, but, nonetheless, it is possible to use the scales to draw comparisons between two states or, for that matter, comparisons between a state and the average scores of a group of states.

Clearly, such comparisons need to be tempered by considerations such as how long the state(s) has been in the SCDP program, the demographics of the state, the size of the state, and the state's educational structure, to name but a few. What is important to note, however, is that the scales permit comparisons with respect to a few, salient attributes; without the scales, the points of comparison would be either so numerous or so ill-defined as to render such comparisons virtually useless.

In addition to being able to make comparisons between states or groups of states, it is also possible to use the scales to make comparisons between the positions of a state or group of states at two or more points in time. In other words, the scales can be used to chart growth, because not only are the scales invariant across states, but it has also been shown that the scales are invariant across time, at least for the years 1978 and 1979. Consequently, the same attribute of the dissemination system will be measured by one of the scales in multiple administrations of the instrument. Thus, the changes that a dissemination system goes through can be described in

terms of a few salient characteristics, rather than with respect to a multitude of details. As with the comparisons between states, utilization of the scales permits a concise description of the evolution of a state's dissemination system.

Having described the scales' potential for making comparisons between states and/or between points in time, three important caveats regarding the utilization and interpretation of the scales must be specified: first, no evidence currently exists that would justify making value judgments such as "good" or "bad" based upon whether a state's position on a scale is "high" or "low"; second, since the scales were developed to be measurement instruments that could be applied to all states, they may be insensitive to dissemination system characteristics that are unique to only a few states; and, third, only occasional evidence has been found to support the interpretation of the scales as being not only hierarchical but also developmental. We discuss each of these caveats more fully.

The first caveat concerns the "high/low" relationship with a judgment of "good/bad"; it must be emphasized that such a relationship cannot be drawn from the data available to this study. Specifically, while the information regarding state dissemination systems with respect to a number of attributes was sufficient to construct a set of scales, value judgments regarding positions on these scales must be based on some external criterion. Such a criterion might ultimately be the degree to which new knowledge is reflected in the improved quality of education in a state. Other criteria might be the extent to which the educators in the state are aware of new knowledge or, for that matter, are aware of and utilize the dissemination systems. In the absence of any external criteria, it is unwise to even begin to speculate how a state's position with respect to the scales should be judged as "good" or

"bad." Given considerations such as geography, size of the school system, and political constraints, a combination of a moderately comprehensive resource base and a moderately comprehensive linkage system might be most appropriate for one state, while for another state, little emphasis on resource base comprehensiveness with a heavy emphasis on linkage might be most appropriate. Furthermore, depending upon the degree to which a dissemination system is developed, it may be that the position of the system with respect to certain scales is emphasized less, while the position of the system with respect to other scales is emphasized more. For example, during the initial stages of system development, rapid growth with respect to comprehensiveness of the resource base might be particularly important, while in a later stage, the size of the resource base might actually shrink as only those resources deemed most essential are retained and monies are reallocated to develop other aspects of the system.

In the absence of any external criteria, a value judgment about a state's position on the scale must necessarily be based on the observer's perceptions as to what constitutes an effective dissemination system. If such an observer is prepared to make such statements as, "the more resources you have, the better" or, "the more coordination you have, the better," then clearly, a high/low position on a scale can be translated into a good/bad judgment. We believe that such an approach to making judgments would be terribly simplistic, however, given the extreme diversity in needs and context that we have encountered in the course of this research. Given the influence that needs and context can have with respect to what sort of dissemination system configuration will be maximally useful in improving instruction, a clear understanding of how these influences operate and the extent to which they are present in a particular state must be taken into

account before any value judgment regarding that state's dissemination system configuration can be rendered.

The second caveat concerns the presence of unique dissemination system characteristics. Since the scales were developed to be applicable across all states, the scales could not include those attributes whose presence or absence in a state was not predictable from the presence or absence of other attributes in that state. While such a lack of predictability may be attributable to the more or less random occurence of that attribute in any particular state, it may also be attributable to the fact that, for a few states, possession of that attribute was particularly important, regardless of what other attributes they might or might not have. For example, a few states obtained the Fugitive Information Data Organizer (FIDO) as part of their resource base, an act which appeared to be completely independent of whether those states had a great many or only a few other resources. Consequently, FIDO was excluded from the comprehensive resource base scale. Thus, for those states subscribing to FIDO, the resource base scale does not completely reflect all of the resources that these states have at their command. Consequently, it must be recognized that while the scales are extremely powerful for purposes of comparison across states and/or across points in time, nonetheless, the sensitivity of the scales to all of the attributes that can constitute a dissemination system is limited. In using the scales to make a comparison between two states, while in one sense the comparison is valid since the scales are measured in the same attributes in both states, in another sense the comparison is not valid if one or both of the states possess a number of attributes that are not included in the scales. It was for this reason that we listed in Chapter 3 all the attributes that were not included in the primary scales. Anybody seeking to use the scales should be

aware of those attributes that were not included and temper any interpretations accordingly.

The third caveat concerns the degree to which each scale can be interpreted as reflecting a developmental pattern. Up to this point in the discussion, the only assumption that has been made about the scales is that they are hierarchical. Given that the scales are hierarchical and that this hierarchy is invariant across both states and points in time, it is appropriate to use the scales to assess differences between the states and to assess growth. Should the scales not only be hierarchical but also developmental, then their interpretive power would be enhanced considerably. In particular, knowledge about where a state is with respect to each of the scales could be used to establish a set of expectations regarding the continued development of the state's dissemination system. Put another way, the indicants in each scale could serve as a series of "mileposts" which would enable us to chart a state's progress in building dissemination capacity.

In Chapter 4, however, we found only occasional evidence to support the interpretation of the scales as being developmental as well as hierarchical. The preceeding statement needs to be tempered, however, by an awareness of the extremely limited amount of data that was available for these analyses. Since information was collected only in 1978 and 1979, we were unable to chart the growth of the Cohort I and Cohort II states during the initial three and two years, respectively, of their participation in the SCBP program. Since the states in the first two Cohorts constituted the bulk of our sample, only minimal information was available regarding dissemination system growth during the initial phase of SCBP participation. With so few observations, the precision with which we could describe the relationship between scale position and years of participation was substantially impaired.

Clearly, when a point in the graphs presented in Figures 4.1 - 4.6 represents only two or three states, the position of that point is far from stable. Even small variations with respect to just one state's score would cause that point to shift position. Furthermore, the data analyses in Chapter 4 suggested that much of the dissemination system's growth might occur during the first two years and that, by years three and four, growth tends to level off.

In addition to the difficulties associated with trying to describe the growth of dissemination systems during the first few years of SCBP participation, our ability to observe and describe dissemination system growth is further impaired by the overpowering impact that contextual characteristics, such as political constraints, can have on the development of a dissemination system. Each of the states in our sample is unique in so many ways that it is extremely difficult to perceive anything that might be common to all states. This problem, which tends to be present in all social science research, is that the variations between individuals or groups of individuals (such as states) tend to be so large that patterns of behavior common to all individuals are obscured. The usual approach to dealing with such a problem is to increase the number of observations, which, of course, was not an option available to us in this study. Consequently, the state and SEA characteristics, especially those characteristics that led certain states to be early participants in the SCBP program, may have completely obscured the relationship between dissemination system development and years of participation in the SCBP program.

Thus, while we were unable to find much support in the available data for a developmental interpretation of the scales, it must be emphasized that this does not rule out the possibility that the scales represent a development

methaphor. It is quite likely, given the limited amount of available data, that the developmental metaphor could have been obscured by the absence of sufficient information describing how dissemination systems grow in the early years of their inception and by the unique characteristics of the states.

Having now considered how the scales can be used to assess differences and to assess growth, and having also considered the three major caveats that must be considered in interpreting and utilizing scale scores, we turn to a discussion of two applications for which we believe the scales are particularly well suited. The first major use of the scale scores would be to permit those associated with a state's dissemination system, as well as those external to it, to assess the system's status. In particular, the scales can be used to establish contexts against which the status of the state's dissemination system configuration can be assessed.

The most obvious of these contexts is one which permits the state to be compared to the dissemination system configurations of all of the other states in this study. This would be what educational researchers refer to as "norm-referenced" interpretation, in which the status of a dissemination system relative to dissemination systems throughout the country would be assessed. Such an assessment could be used to indicate aspects of the dissemination system where further development is needed, as well as those aspects in which a sufficiently high level has already been achieved. Such an assessment, of course, would need to be tempered by an awareness of whether the dissemination system was still in a fledgling state or more advanced stage of development, as well as a realization that certain components of the dissemination system might not be reflected in the scale scores. In addition, such comparisons would have to be tempered by the unique political, demographic, and educationa considerations that may exist in a state.

Taking all of these considerations into account, however, a norm-referenced interpretation could be useful in indicating what aspects of a dissemination system might be targeted for further development.

This same information could also be used by NIE in assessing the proposed emphases contained in the SCBP applications. Given a profile of a state's dissemination system with respect to each of the scales, a state which proposes to continue to spend SCBP funds on development of its resource base and yet which already ranks well above the national average with respect to its resource base might be questioned as to the appropriateness of such a priority. Of course, the political, demographic, and educational constraints under which the state must operate should all be taken into account in making such a judgment. For instance, a large state with a sparse population might find it much more appropriate to invest in the continued development than might a small state with a very dense population.

To summarize, then, the scales can be used by both the states and NIE to assess the status of a dissemination system configuration relative to the dissemination systems of other states. While such an assessment must take into account certain considerations that will be unique to a state, nonetheless, the scales can provide some guidelines as to where additional development might be needed or not needed. Clearly, the usefulness of the scales will be enhanced with the inclusion of more information in the decision-making process, such as the scores of the dissemination system at some previous time.

The second major use of the scale scores would be to measure change as well as status. Furthermore, change can be measured not just over two points in time, but across several points in time. From the state perspective, the ability to assess change using the scales has several important applications.

To begin with, objectives can be stated at the beginning of the year with regard to how the dissemination system configuration will be modified, and the degree to which these objectives have been met can then be assessed at the end of the year by comparing the change in the scale scores. For instance, if a state has as one of its objectives increasing the accessibility of the dissemination system to users, the amount of change in the score on the coordination of the resource base scale can be used, at least partially, to assess the extent to which this objective has been met. By the same token, changes with respect to the institutionalization scale can be used to partially assess the degree to which the dissemination system functions are being underwritten by the state as opposed to requiring federal support. If data on the scales are available for several years, then this retrospective information could potentially be used to establish objectives and expectations about what areas or aspects of the dissemination system should be targeted for further development.

As might be expected, virtually the same sorts of uses apply to NIE as well. In the course of monitoring the SCBP project, if the project application targets some aspect of the dissemination system for further development and no change occurs with respect to the appropriate scale score, then this might be taken as an indication by NIE staff that the operation of the program with respect to the objective should be scrutinized further. It may well be that the political, demographic, and educational constraints in that state acted in some fashion to render meeting the objective impossible. Nonetheless, the scale scores do provide a convenient and concise way of identifying areas where the states are having difficulty meeting their objectives. In addition, if multiple data points are available, the NIE staff can take a retrospective view of dissemination system development in a state and based upon that review have certain expectations about what sorts

of emphases should be in the next project application. We are not suggesting that these expectations should necessarily be based on some sort of developmental model, since as was stated in a previous chapter, no evidence has yet been found to support the evidence of any sort of developmental paradigm. However, NIE monitors who are familiar with the states and their unique characteristics could combine that knowledge with the profile and the knowledge of how the SCBP program has been operating to arrive at some expectations regarding what future directions the SCBP program should take.

SUMMARY

Even with the constraints that the caveats discussed in this chapter place upon their use, the scales can be a powerful interpretative tool in assessing a state's dissemination system configuration. We do not believe, however, that the construction and application of scales such as those that have been described in this substudy are necessarily limited to dissemination system configurations. Given that the relevent attributes underlying an organization can be identified and described, then it is likely that a set of indicants can be generated which will describe each of the attributes and a scale can be constructed. As this study has shown, it is not unreasonable to expect that many of the indicants will form a hierarchy, and hence can be legitimately considered to constitute a scale. Such a scale can be used to not only provide a succinct description of an organization, but can also be used as a basis for comparing organizations. Should the hierarchy of the indicants remain invariant across repeated administrations of the indicants, then the scale can be used to measure organizational growth. Finally, if that organizational growth can be shown to follow a consistent pattern, then the scale can be used to not only measure growth to date, but also to

establish expectations regarding future growth. While this last possibility was not present in the dissemination system scales, the usefulness of the scales in making comparisons and measuring growth is capable of providing individuals both within and external to the organization with a powerful tool in assessing the status of their organization. Such assessments can be used as a valuable source of input to the decision-making process regarding what sorts of activities related to the organization should be undertaken or terminated.

REFERENCES

The NTS Research Corporation reports for this study all have the general title Building Capacity for the Improvement of Educational Practice. The complete set of volume is as follows:

Volume I — Final Report: Building Capacity for the Improvement of Education, An Evaluation of NIE's State Dissemination Grants Program (April 1981)

Volume II — 1979 State Abstracts: State Dissemination Efforts (April 1980)

Volume III — A Study of Linker Agent Activities and Roles (April 1981)

Volume IV — A Study of the Development of Scales Measuring Dissemination Capacity (April 1981)

Volume V — Executive Summary (April 1981)

For further information contact:

NTS RESEARCH CORPORATION

2634 Chapel Hill Boulevard
Durham, NC 27707
(919) 493-2451

Eugene C. Royster
Principal Investigator

Doren L. Madey
Project Director

NATIONAL INSTITUTE OF EDUCATION

Program on Dissemination and Improvement of Practice
Research and Educational Practice Unit
(202) 254-6050

Michael B. Kane
NIE Assistant Director

John C. Egermeier
Project Officer